

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
**Федеральное государственное автономное образовательное учреждение  
высшего образования**

Приложение 4  
к ОПОП ВО 01.04.04 ПРИКЛАДНАЯ МАТЕМАТИКА,  
профиль "Математические методы в искусственном интеллекте  
и анализе данных"

**Рабочая программа дисциплины (модуля)**

**Введение в искусственные нейронные сети**

Закреплена за подразделением

Кафедра инженерной кибернетики

Направление подготовки

01.04.04 ПРИКЛАДНАЯ МАТЕМАТИКА

Профиль

Математические методы в искусственном интеллекте и анализе данных

Квалификация

**Магистр**

Форма обучения

**очная**

Общая трудоемкость

**4 ЗЕТ**

Часов по учебному плану

144

Формы контроля в семестрах:

в том числе:

зачет с оценкой 1

аудиторные занятия

34

самостоятельная работа

110

**Распределение часов дисциплины по семестрам**

Семестр (<Курс>.<Семестр на курсе>)	1 (1.1)		Итого	
Неделя	18			
Вид занятий	уп	рп	уп	рп
Лекции	17	17	17	17
Практические	17	17	17	17
Итого ауд.	34	34	34	34
Контактная работа	34	34	34	34
Сам. работа	110	110	110	110
Итого	144	144	144	144

Программу составил(и):

*д.т.н., проф., Садеков Ринат Наилевич*

Рабочая программа

**Введение в искусственные нейронные сети**

Разработана в соответствии с ОС ВО:

Самостоятельно устанавливаемый образовательный стандарт высшего образования - магистратура Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский технологический университет «МИСИС» по направлению подготовки 01.04.04 ПРИКЛАДНАЯ МАТЕМАТИКА (приказ от 02.04.2021 г. № 119 о.в.)

Составлена на основании учебного плана:

01.04.04 Прикладная математика, 01.04.04-МПИМ-24-1.plx Математические методы в искусственном интеллекте и анализе данных, утвержденного Ученым советом НИТУ МИСИС в составе соответствующей ОПОП ВО 22.06.2023, протокол № 5- 23

Утверждена в составе ОПОП ВО:

01.04.04 Прикладная математика, Математические методы в искусственном интеллекте и анализе данных, утвержденной Ученым советом НИТУ МИСИС 22.06.2023, протокол № 5-23

Рабочая программа одобрена на заседании

**Кафедра инженерной кибернетики**

Протокол от 20.06.2023 г., №11

Руководитель подразделения Ефимов А.Р.

1. ЦЕЛИ ОСВОЕНИЯ	
1.1	Дать систематический обзор современных методов анализа информации с использования нейросетевых моделей, изучить и освоить принципы и современные технологии используемые для построения и использования нейронных сетей, рассмотреть перспективные направления развития моделей и методов.
1.2	По результатам освоения дисциплины студенты научатся анализировать прикладные задачи, выбирать методы их решения и создавать программное обеспечение с использованием известных фреймворков и технологий

2. МЕСТО В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ	
Блок ОП:	Б1.О
2.1	Требования к предварительной подготовке обучающегося:
2.2	Дисциплины (модули) и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:
2.2.1	Алгоритмизация и программирование
2.2.2	Искусственный интеллект в задачах распознавания образов
2.2.3	Методы машинного обучения
2.2.4	Научно-исследовательская практика
2.2.5	Производственная практика
2.2.6	Блокчейн - технологии
2.2.7	Интеллектуальные автономные и мультиагентные системы
2.2.8	Машинное обучение и методология DevOps при разработке систем искусственного интеллекта
2.2.9	Научно-исследовательская работа
2.2.10	Системный подход и генерация знаний в инновациях
2.2.11	Современные устройства центров обработки больших данных
2.2.12	Экспертные и рекомендательные, информационно-аналитические системы
2.2.13	Методы искусственного интеллекта в робототехнических системах
2.2.14	Подготовка к процедуре защиты и защита выпускной квалификационной работы
2.2.15	Преддипломная практика
2.2.16	Философия, методология и современные тренды искусственного интеллекта как науки

3. РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫЕ С ФОРМИРУЕМЫМИ КОМПЕТЕНЦИЯМИ	
<b>ОПК-2: Способен разрабатывать и развивать математические методы моделирования объектов, процессов и систем в области профессиональной деятельности, осуществлять моделирование, анализ и эксперименты в целях проведения детального исследования для решения сложных задач</b>	
<b>Знать:</b>	
ОПК-2-33 Методы градиентного спуска и обратного распространения ошибки	
ОПК-2-32 Назначение основных слоев, функции активации, функции ошибок нейронных сетей	
ОПК-2-31 Основные архитектуры нейронных сетей используемые для решения задач по обработке текстовой, визуальной и другой информации	
<b>ОПК-1: Способен обобщать и критически оценивать опыт и результаты научных исследований в области прикладной математики, на основе знаний фундаментальных наук, в междисциплинарных областях</b>	
<b>Знать:</b>	
ОПК-1-31 Знать основные источники и методы поиска научной информации по направлению нейронные сети	
ОПК-1-32 Знать основной круг проблем (задач), встречающихся в нейронных сетях, и основные способы (методы, алгоритмы) их решения	
<b>ОПК-2: Способен разрабатывать и развивать математические методы моделирования объектов, процессов и систем в области профессиональной деятельности, осуществлять моделирование, анализ и эксперименты в целях проведения детального исследования для решения сложных задач</b>	
<b>Уметь:</b>	
ОПК-2-У1 Подбирать архитектуру нейронных сетей под любой класс задач (анализ текстовой, визуальной и др. инф)	
ОПК-2-У2 Подбирать методы обучения нейронных сетей, функции активации и ошибок для любого класса задач	

<b>ОПК-1: Способен обобщать и критически оценивать опыт и результаты научных исследований в области прикладной математики, на основе знаний фундаментальных наук, в междисциплинарных областях</b>								
<b>Уметь:</b>								
ОПК-1-У1 Собирать, отбирать и использовать необходимые для обучения нейронных сетей датасеты								
ОПК-1-У2 Анализировать, систематизировать, усваивать и оценивать передовой опыт по использованию нейронных сетей изложенный в научных статьях								
ОПК-1-У3 Находить наиболее эффективные (методы) решения основных типов проблем (задач)								
ОПК-1-У4 Находить в открытом доступе предобученные модели нейронных сетей и понимать границы их использования и потенциальный эффект								
<b>ОПК-2: Способен разрабатывать и развивать математические методы моделирования объектов, процессов и систем в области профессиональной деятельности, осуществлять моделирование, анализ и эксперименты в целях проведения детального исследования для решения сложных задач</b>								
<b>Владеть:</b>								
ОПК-2-В2 Современными фреймворками используемыми для работы с нейронными сетями в части создания нейронных сетей заданной архитектуры								
ОПК-2-В1 Современными фреймворками используемыми для работы с нейронными сетями в части использования предобученных моделей, инференса и обучения нейронных сетей								
<b>ОПК-1: Способен обобщать и критически оценивать опыт и результаты научных исследований в области прикладной математики, на основе знаний фундаментальных наук, в междисциплинарных областях</b>								
<b>Владеть:</b>								
ОПК-1-В1 Информационно-аналитическими инструментами для поиска информации освещающей современные достижения в области нейронных сетей								
ОПК-1-В2 Навыками по регистрации на открытых интернет площадках (сайтах) для получения информации отражающей текущее состояние дел в предметной области								

#### 4. СТРУКТУРА И СОДЕРЖАНИЕ

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Формируемые индикаторы компетенций	Литература и эл. ресурсы	Примечание	КМ	Выполняемые работы
	<b>Раздел 1. Общие сведения о нейронных сетях</b>							
1.1	Введение /Лек/	1	2	ОПК-1-31 ОПК-1-32 ОПК-1-У2 ОПК-1-В1	Л1.2 Л1.3 Л1.4 Э1		КМ1	Р1
1.2	История развития нейронных сетей /Лек/	1	1	ОПК-1-31 ОПК-1-32 ОПК-1-У2 ОПК-1-В1 ОПК-2-32	Л1.1 Л1.2 Л1.3 Л1.4 Э2		КМ2	Р1
1.3	Модель перцептрона /Лек/	1	1	ОПК-1-31 ОПК-2-31 ОПК-2-32	Л1.1 Л1.2 Л1.3 Л1.4 Э3		КМ5	Р2
1.4	Метод градиентного спуска /Лек/	1	1	ОПК-2-32 ОПК-2-33	Л1.2 Л1.3 Л1.4 Э5		КМ3	Р3
1.5	Метод обратного распространения ошибки /Лек/	1	1	ОПК-2-32 ОПК-2-33	Л1.2 Л1.3 Л1.4 Э5		КМ3	Р4
1.6	Методы оптимизации /Лек/	1	1	ОПК-2-32 ОПК-2-33	Л1.2 Л1.3 Л1.4 Э2		КМ5	Р4
1.7	Функции активации /Лек/	1	1	ОПК-2-32 ОПК-2-33	Л1.2 Л1.3 Л1.4 Э8		КМ6	Р4
1.8	Инициализация весов, нормализация и регуляризация /Лек/	1	1	ОПК-1-У1 ОПК-2-31 ОПК-2-32	Л1.2 Л1.3 Л1.4 Э5		КМ6	Р5
1.9	Аугментация данных /Лек/	1	2	ОПК-1-У1 ОПК-1-В1 ОПК-2-32	Л1.2 Л1.3 Л1.4 Э2		КМ2	Р4

1.10	Сравнительная оценка методов оптимизации нейронных сетей /Ср/	1	4	ОПК-1-У1 ОПК-1-У4 ОПК-1-В2	Л1.2 Л1.3 Л1.4 Э4		КМ1,КМ 2	Р6
1.11	Сравнительная оценка функций активации нейронных сетей /Ср/	1	4	ОПК-1-У1 ОПК-1-У3 ОПК-1-У4 ОПК-1-В2 ОПК-2-32 ОПК -2-33 ОПК-2- У2	Л1.2 Л1.3 Л1.4 Э5		КМ2,КМ 4	Р2
1.12	Сравнительная оценка методов инициализации, нормализации и регуляризации /Ср/	1	8	ОПК-1-У1 ОПК-1-У3 ОПК-1-У4 ОПК-1-В2 ОПК-2-32 ОПК -2-33 ОПК-2- У2	Л1.2 Л1.3 Л1.4 Э2		КМ2,КМ 3	Р5
1.13	Датасеты используемые для решения различных типовых задач /Ср/	1	16	ОПК-1-У1	Л1.2 Л1.3 Л1.4 Э2		КМ5,КМ 4,КМ7	Р3
1.14	Сравнительная оценка подходов к аугментации данных /Ср/	1	8	ОПК-1-31 ОПК -1-У1 ОПК-1- В2	Л1.2 Л1.3 Л1.4 Э3		КМ8,КМ 5	Р8
	<b>Раздел 2. Популярные архитектуры нейронных сетей</b>							
2.1	Конволюционные нейронные сети /Лек/	1	2	ОПК-1-31 ОПК -1-У2 ОПК-1- В1 ОПК-2-У1 ОПК-2-В1	Л1.2 Л1.3 Л1.4	Литература для всех видов занятий этого раздела одинаковая	КМ1	Р1
2.2	Основы Pytorch /Пр/	1	2	ОПК-1-31 ОПК -1-У3 ОПК-2- У1 ОПК-2-В1	Л1.2 Л1.3 Л1.4 Э3		КМ1	Р1
2.3	Начало работы с глубоким обучением ("Getting Started with Deep Learning") /Пр/	1	3	ОПК-1-У1 ОПК-1-У3 ОПК-1-У4 ОПК-1-В1 ОПК-1-В2 ОПК-2-33 ОПК -2-У1	Л1.2 Л1.3 Л1.4 Э4		КМ2	Р2
2.4	Сети для сегментации ("Getting Started with Image Segmentation") /Пр/	1	2	ОПК-1-У1 ОПК-1-У2 ОПК-1-У4 ОПК-1-В1 ОПК-1-В2 ОПК-2-У1	Л1.2 Л1.3 Л1.4 Э5		КМ3	Р3
2.5	Классификация изображений с фреймворком DIGITS ("Image Classification with DIGITS") /Пр/	1	2	ОПК-1-У1 ОПК-1-У4 ОПК-1-В1 ОПК-2-У1 ОПК-2-У2 ОПК-2-В1	Л1.2 Л1.3 Л1.4 Э1		КМ4	Р4
2.6	Использование фреймворка horovod ("Deep Learning at Scale with Horovod") /Пр/	1	2	ОПК-1-У1 ОПК-1-У4 ОПК-2-В1	Л1.2 Л1.3 Л1.4 Э6		КМ1	Р1
2.7	Оптимизация сетей с использованием фреймворка tensorRT ("Optimization and Deployment of TensorFlow Models with TensorRT") /Пр/	1	2	ОПК-1-У1 ОПК-1-У4 ОПК-1-В1 ОПК-2-У2 ОПК-2-В1 ОПК-2-В2	Л1.2 Л1.3 Л1.4 Э7		КМ6	Р6

2.8	Использование контейнеров ("High-Performance Computing with Containers") /Пр/	1	2	ОПК-1-У1 ОПК-1-У4 ОПК-1-В1 ОПК-2-В2	Л1.2 Л1.3 Э8		КМ7	Р7
2.9	Создание систем видеоаналитики ("AI Workflows for Intelligent Video Analytics with DeepStream") /Пр/	1	2	ОПК-1-У1 ОПК-1-У4 ОПК-1-В1 ОПК-2-У2 ОПК-2-В1 ОПК-2-В2	Э9		КМ8	Р8
2.10	Основные возможности AWS облака для реализации технологий нейронных сетей /Ср/	1	8	ОПК-2-31 ОПК-2-32 ОПК-2-33 ОПК-2-У1 ОПК-2-У2 ОПК-2-В1 ОПК-2-В2 ОПК-1-32 ОПК-1-31 ОПК-1-У1 ОПК-1-У2 ОПК-1-У3	Э2		КМ1	Р4
2.11	Возможности pytorch для обучения нейронных сетей на мультигиг платформах /Ср/	1	8	ОПК-1-В2 ОПК-2-33 ОПК-2-У1 ОПК-2-У2	Э3		КМ1	Р6
2.12	Возможности gstream для реализации систем видеоаналитики для работы в реальном времени /Ср/	1	10	ОПК-1-У4 ОПК-1-У3 ОПК-1-У2 ОПК-1-У1	Э9		КМ1	Р6
2.13	Популярные зоопарки нейронных сетей для использования в собственных алгоритмах /Ср/	1	12	ОПК-1-31 ОПК-1-В2 ОПК-2-У1	Э3 Э4		КМ7	Р2
2.14	Настройка инфраструктуры для обучения нейронных сетей /Ср/	1	6	ОПК-2-31 ОПК-2-У1 ОПК-2-В1 ОПК-2-У2	Э2		КМ6	Р2
2.15	Подготовка к сдаче и защите практических занятий /Ср/	1	10	ОПК-2-31 ОПК-2-32 ОПК-2-33 ОПК-2-У1 ОПК-1-В1 ОПК-1-В2 ОПК-1-У4	Э1 Э2 Э4 Э5 Э6 Э7 Э8 Э9		КМ6	Р6
	<b>Раздел 3. Перспективы развития нейросетевых подходов</b>							
3.1	Архитектура GAN сетей /Лек/	1	1	ОПК-2-31 ОПК-2-32 ОПК-2-У1 ОПК-1-31 ОПК-1-32 ОПК-1-У3	Л1.2 Л1.3 Л1.4	Литература для всех видов занятий этого раздела одинаковая	КМ1	Р1
3.2	Архитектура RNN сетей /Лек/	1	1	ОПК-1-31 ОПК-1-32 ОПК-1-У2 ОПК-2-31 ОПК-2-33			КМ2	Р2
3.3	Перспективы развития нейронных сетей /Лек/	1	2	ОПК-1-31 ОПК-1-32 ОПК-1-У2 ОПК-2-31			КМ3	Р3

3.4	Перспективы развития GAN сетей /Ср/	1	4	ОПК-1-31 ОПК-1-32 ОПК-1-У2 ОПК-2-31 ОПК-2-33			КМ4	Р4
3.5	Перспективы развития RNN сетей /Ср/	1	4	ОПК-1-31 ОПК-1-32 ОПК-1-У2 ОПК-2-31 ОПК-2-33			КМ1	Р3
3.6	Существующие и перспективные архитектуры трансформеров /Ср/	1	8	ОПК-1-31 ОПК-1-32 ОПК-1-У2			КМ8	Р8

## 5. ФОНД ОЦЕНОЧНЫХ МАТЕРИАЛОВ

### 5.1. Контрольные мероприятия (контрольная работа, тест, коллоквиум, экзамен и т.п), вопросы для самостоятельной подготовки

Код КМ	Контрольное мероприятие	Проверяемые индикаторы компетенций	Вопросы для подготовки
КМ1	Тест №1 "Основы Pytorch"	ОПК-2-31;ОПК-2- У1	<p>Что такое PyTorch?</p> <p>В чем разница между PyTorch и другими библиотеками машинного обучения?</p> <p>Что из перечисленного не является особенностью PyTorch?</p> <p>Как работает автоматическое дифференцирование в PyTorch?</p> <p>Какой модуль PyTorch используется для создания графического процессора (GPU)?</p> <p>Какие типы слоев поддерживаются в PyTorch?</p> <p>Как создать модель в PyTorch для классификации изображений?</p> <p>Как обучить модель в PyTorch с использованием GPU?</p>
КМ2	Тест №2 Начало работы с глубоким обучением ("Getting Started with Deep Learning")	ОПК-2-33;ОПК-2-У2;ОПК-1-В1	<p>Каковы основные компоненты архитектуры глубокого обучения?</p> <p>Как вы определяете оптимальные гиперпараметры для модели глубокого обучения?</p> <p>В чем преимущества использования сверточных нейронных сетей (CNN) в задачах компьютерного зрения?</p> <p>Что такое регуляризация и как она используется для предотвращения переобучения в глубоком обучении?</p> <p>Опишите различные типы функций активации, используемых в глубоком обучении, и объясните, как они работают.</p> <p>Что такое метод обратного распространения ошибки и как он используется для обучения нейронных сетей?</p> <p>В чем заключается разница между обучением с учителем и без учителя в контексте глубокого обучения?</p> <p>Объясните, что такое пакетная нормализация и почему она важна для эффективного обучения глубоких нейронных сетей.</p> <p>Чем отличаются рекуррентные нейронные сети от обычных многослойных перцептронов?</p> <p>Что такое трансферное обучение и как оно может быть использовано для ускорения разработки моделей глубокого обучения?</p>

КМ3	Тест №3 Сети для сегментации ("Getting Started with Image Segmentation")	ОПК-2-32;ОПК-2-У2;ОПК-1-У3	<p>Дайте определение задаче сегментации изображений.</p> <p>Что такое U-Net и каковы его основные характеристики?</p> <p>Как вы можете использовать сверточные нейронные сети (CNN) для сегментации изображений?</p> <p>В чем преимущество использования полностью сверточных сетей (FCN) для сегментации?</p> <p>Что такое генеративно-сопоставительные сети (GAN) и как они используются для сегментации изображений?</p> <p>В чем отличие использования семантической сегментации от instance segmentation?</p> <p>Что такое проточная архитектура и как она применяется для сегментации изображения?</p> <p>Что такое пиксельная классификация и как она связана с сегментацией изображений?</p> <p>Как вы выбираете метрику для оценки качества сегментации изображения?</p> <p>В каких приложениях может быть полезна сегментация изображений?</p>
КМ4	Тест №4 Классификация изображений с фреймворком DIGITS ("Image Classification with DIGITS")	ОПК-1-В2;ОПК-1-В1	<p>Опишите архитектуру фреймворка для классификации изображений DIGITS.</p> <p>Каковы основные этапы процесса классификации изображений с использованием DIGITS?</p> <p>Какие модели CNN поддерживаются фреймворком DIGITS для классификации изображений?</p> <p>Какие функции доступны для предварительной обработки изображений в DIGITS?</p> <p>Как осуществляется оптимизация гиперпараметров в DIGITS для повышения точности классификации изображений?</p> <p>Как выполняется оценка качества модели классификации изображений в DIGITS?</p> <p>Какие метрики используются для оценки точности модели классификации изображений?</p> <p>Опишите процесс сохранения и загрузки модели классификации изображений, созданной с помощью DIGITS.</p> <p>Какие методы визуализации доступны в DIGITS для анализа и интерпретации результатов классификации изображений?</p> <p>Какие расширения и дополнительные возможности предоставляет фреймворк DIGITS по сравнению с традиционным ручным кодированием и настройкой модели классификации изображений?</p>
КМ5	Тест №5 Использование фреймворка horovod ("Deep Learning at Scale with Horovod")	ОПК-2-У1;ОПК-2-31;ОПК-1-У3	<p>Какова цель фреймворка Horovod?</p> <p>Какие основные преимущества использования Horovod для масштабирования глубокого обучения?</p> <p>Какие библиотеки и технологии использует Horovod?</p> <p>Опишите архитектуру Horovod и как она работает.</p> <p>Каковы ограничения использования Horovod в некоторых сценариях?</p> <p>Как Horovod обеспечивает эффективное использование графических процессоров (GPU) в кластере?</p> <p>Какую роль играет алгоритм AllReduce в Horovod?</p> <p>В чем заключаются преимущества использования библиотеки TensorFlow с Horovod по сравнению с PyTorch?</p> <p>Опишите процесс развертывания модели с использованием Horovod на кластере с несколькими узлами.</p> <p>Как можно оптимизировать производительность модели, обученной с помощью Horovod, с помощью гиперпараметров и конфигурации?</p>



КМ6	Тест №6 Оптимизация сетей с использованием фреймворка tensorRT ("Optimization and Deployment of TensorFlow Models with TensorRT")	ОПК-1-B2;ОПК-1-У4;ОПК-1-B1	<p>Какова основная цель фреймворка TensorRT?</p> <p>Опишите основные преимущества использования TensorRT для оптимизации и развертывания сетей TensorFlow.</p> <p>Сравните производительность TensorRT с другими фреймворками оптимизации моделей машинного обучения, такими как Caffe2 и ONNX.</p> <p>Какая архитектура используется в TensorRT для обеспечения высокой производительности и эффективности при инференсе моделей TensorFlow?</p> <p>Какие оптимизации производительности доступны в TensorRT и как они влияют на инференс моделей TensorFlow?</p> <p>Опишите процесс преобразования модели TensorFlow в формат, поддерживаемый TensorRT.</p> <p>Каким образом TensorRT обеспечивает оптимальное использование графических процессоров и других аппаратных ресурсов для инференса моделей TensorFlow?</p> <p>Какие дополнительные функции предоставляет TensorRT по сравнению с базовой реализацией инференса TensorFlow?</p> <p>В чем состоят ограничения использования TensorRT в некоторых сценариях развертывания моделей TensorFlow, таких как обучение или онлайн-инференс?</p> <p>Каков процесс развертывания оптимизированной модели TensorRT на целевом устройстве или в облачной инфраструктуре?</p>
КМ7	Тест №7 Использование контейнеров ("HighPerformance Computing with Containers")	ОПК-1-B2;ОПК-1-У4;ОПК-2-B2	<p>Какие преимущества предоставляет использование контейнеров для развертывания и масштабирования приложений машинного обучения и глубокого обучения?</p> <p>Сравните и сопоставьте различные системы оркестровки контейнеров, такие как Docker Swarm, Kubernetes и Mesos, с точки зрения их функциональности, производительности и стоимости.</p> <p>Опишите жизненный цикл разработки и развертывания приложения машинного обучения или глубокого обучения с использованием контейнеров.</p> <p>Каковы наилучшие практики для настройки и оптимизации производительности контейнеров для приложений машинного обучения и глубокого обучения?</p> <p>В чем состоит роль стандартов, таких как Docker Hub, для хранения и обмена образами контейнеров в экосистеме разработки машинного обучения и глубокого обучения?</p> <p>Каковы недостатки использования контейнеров для развертывания приложений машинного обучения и глубокого обучения, такие как сложность управления, стоимость и ограничения безопасности?</p> <p>Опишите процесс миграции существующего приложения машинного обучения или глубокого обучения на использование контейнеров для обеспечения масштабируемости, повторного использования кода и переносимости.</p> <p>Сравните и сопоставьте подходы к управлению ресурсами, планированию и оркестрации для контейнеров, используемые в различных системах, таких как Kubernetes, Mesos и Docker Swarm.</p> <p>Как контейнеры облегчают развертывание моделей машинного обучения и глубокого обучения в производство?</p>

КМ8	Тест №8 Создание систем видеоаналитики ("AI Workflows for Intelligent Video Analytics with DeepStream")	ОПК-2-31;ОПК-2-32;ОПК-2-У1;ОПК-2-33;ОПК-2-У2;ОПК-2-В1	<p>В чём заключается роль систем видеоаналитики в современных приложениях и инфраструктурах?</p> <p>Дайте обзор фреймворка DeepStream и опишите его основные возможности.</p> <p>Какие задачи компьютерного зрения могут быть решены с использованием DeepStream?</p> <p>Сравните фреймворк DeepStream с другими решениями для видеоаналитики, такими как OpenVINO и Darknet.</p> <p>Каким образом DeepStream интегрируется с платформами и системами видеонаблюдения, такими как Intel® OpenVINO™ и ONVIF?</p> <p>Охарактеризуйте процесс обработки видео с использованием DeepStream, включая этапы обнаружения объектов, классификации и отслеживания.</p> <p>Какова роль сверточных нейронных сетей (CNN) и глубокого обучения (DL) в алгоритмах видеоаналитики на базе DeepStream?</p> <p>Как обеспечивается масштабируемость и гибкость систем видеоаналитики с использованием DeepStream в распределенных и облачных инфраструктурах?</p> <p>Каковы преимущества использования DeepStream для развертывания систем видеоаналитики по сравнению с традиционными решениями?</p> <p>Опишите процесс внедрения и интеграции DeepStream в существующие системы видеонаблюдения и машинного зрения.</p>
-----	---	---	--

### 5.2. Перечень работ, выполняемых по дисциплине (Курсовая работа, Курсовой проект, РГР, Реферат, ЛР, ПР и т.п.)

Код работы	Название работы	Проверяемые индикаторы компетенций	Содержание работы
P1	Введение в Pytorch	ОПК-2-31;ОПК-2-32	Знакомство с простейшими нейронными сетями и их реализация на языке программирования python и библиотеки pytorch
P2	Image Classification with DIGITS	ОПК-2-33;ОПК-2- У1	Использование Фреймворка DIGITS для обучения нейронных сетей в задаче классификации и детектирования
P3	Getting Started with Deep Learning	ОПК-2-У2;ОПК-2-В1	Использование основных архитектур нейронных сетей для решения задачи распознавания объекта, классификации, анализа текста
P4	Getting Started with Image Segmentation	ОПК-2-В2;ОПК-1- 31	Использование сегментационной сети для анализа медицинских МРТ снимков
P5	Deep Learning at Scale with Horovod	ОПК-1-32;ОПК-1- У1	Использование глубоких сетей на серверах с несколькими GPU
P6	Optimization and Deployment of TensorFlow Models with TensorRT	ОПК-1-У2;ОПК-1-У3	Исследование возможностей фреймворка TensorRT для ускорение работы нейронной сети
P7	High-Performance Computing with Containers	ОПК-1-У4;ОПК-1-В1;ОПК-1-В2	Использование контейнеров для виртуализации решений
P8	AI Workflows for Intelligent Video Analytics with DeepStream	ОПК-2-31;ОПК-2-У1;ОПК-1-В2;ОПК-1-В1	Обработка потокового видео с использованием нейронных сетей

### 5.3. Оценочные материалы, используемые для экзамена (описание билетов, тестов и т.п.)

Для подготовки вопросов к зачету используются материалы тестов (всего около 150 вопросов):

Что такое PyTorch?

В чем разница между PyTorch и другими библиотеками машинного обучения?

Что из перечисленного не является особенностью PyTorch?

Как работает автоматическое дифференцирование в PyTorch?

Какой модуль PyTorch используется для создания графического процессора (GPU)?

Какие типы слоев поддерживаются в PyTorch?

Как создать модель в PyTorch для классификации изображений?

Как обучить модель в PyTorch с использованием GPU?

Каковы основные компоненты архитектуры глубокого обучения?

Как вы определяете оптимальные гиперпараметры для модели глубокого обучения?

В чем преимущества использования сверточных нейронных сетей (CNN) в задачах компьютерного зрения?

Что такое регуляризация и как она используется для предотвращения переобучения в глубоком обучении?

Опишите различные типы функций активации, используемых в глубоком обучении, и объясните, как они работают.

Что такое метод обратного распространения ошибки и как он используется для обучения нейронных сетей?

В чем заключается разница между обучением с учителем и без учителя в контексте глубокого обучения?

Объясните, что такое пакетная нормализация и почему она важна для эффективного обучения глубоких нейронных сетей.

Чем отличаются рекуррентные нейронные сети от обычных многослойных перцептронов?

Что такое трансферное обучение и как оно может быть использовано для ускорения разработки моделей глубокого обучения?

Что такое U-Net и каковы его основные характеристики?

Как вы можете использовать сверточные нейронные сети (CNN) для сегментации изображений?

В чем преимущество использования полностью сверточных сетей (FCN) для сегментации?

Что такое генеративно-состязательные сети (GAN) и как они используются для сегментации изображений?

В чем отличие использования семантической сегментации от instance segmentation?

Что такое проточная архитектура и как она применяется для сегментации изображений?

Что такое пиксельная классификация и как она связана с сегментацией изображений?

Как вы выбираете метрику для оценки качества сегментации изображения?

В каких приложениях может быть полезна сегментация изображений?

Опишите архитектуру фреймворка для классификации изображений DIGITS.

Каковы основные этапы процесса классификации изображений с использованием DIGITS?

Какие модели CNN поддерживаются фреймворком DIGITS для классификации изображений?

Какие функции доступны для предварительной обработки изображений в DIGITS?

Как осуществляется оптимизация гиперпараметров в DIGITS для повышения точности классификации изображений?

Как выполняется оценка качества модели классификации изображений в DIGITS?

Какие метрики используются для оценки точности модели классификации изображений?

Опишите процесс сохранения и загрузки модели классификации изображений, созданной с помощью DIGITS.

Какие методы визуализации доступны в DIGITS для анализа и интерпретации результатов классификации изображений?

Какие расширения и дополнительные возможности предоставляет фреймворк DIGITS по сравнению с традиционным ручным кодированием и настройкой модели классификации изображений?

Какова цель фреймворка Horovod?

Какие основные преимущества использования Horovod для масштабирования глубокого обучения?

Какие библиотеки и технологии использует Horovod?

Опишите архитектуру Horovod и как она работает.

Каковы ограничения использования Horovod в некоторых сценариях?

Как Horovod обеспечивает эффективное использование графических процессоров (GPU) в кластере?

Какую роль играет алгоритм AllReduce в Horovod?

В чем заключаются преимущества использования библиотеки TensorFlow с Horovod по сравнению с PyTorch?

Опишите процесс развертывания модели с использованием Horovod на кластере с несколькими узлами.

Как можно оптимизировать производительность модели, обученной с помощью Horovod, с помощью гиперпараметров и конфигурации?

Какова основная цель фреймворка TensorRT?

Опишите основные преимущества использования TensorRT для оптимизации и развертывания сетей TensorFlow.

Сравните производительность TensorRT с другими фреймворками оптимизации моделей машинного обучения, такими как Caffe2 и ONNX.

Какая архитектура используется в TensorRT для обеспечения высокой производительности и эффективности при инференсе моделей TensorFlow?

Какие оптимизации производительности доступны в TensorRT и как они влияют на инференс моделей TensorFlow?

Опишите процесс преобразования модели TensorFlow в формат, поддерживаемый TensorRT.

Каким образом TensorRT обеспечивает оптимальное использование графических процессоров и других аппаратных ресурсов для инференса моделей TensorFlow?

Какие дополнительные функции предоставляет TensorRT по сравнению с базовой реализацией инференса TensorFlow?

В чем состоят ограничения использования TensorRT в некоторых сценариях развертывания моделей TensorFlow, таких как обучение или онлайн-инференс?

Каков процесс развертывания оптимизированной модели TensorRT на целевом устройстве или в облачной инфраструктуре?

Какие преимущества предоставляет использование контейнеров для развертывания и масштабирования приложений машинного обучения и глубокого обучения?

Сравните и сопоставьте различные системы оркестровки контейнеров, такие как Docker Swarm, Kubernetes и Mesos, с точки зрения их функциональности, производительности и стоимости.

Опишите жизненный цикл разработки и развертывания приложения машинного обучения или глубокого обучения с использованием контейнеров.

Каковы наилучшие практики для настройки и оптимизации производительности контейнеров для приложений машинного обучения и глубокого обучения?

В чем состоит роль стандартов, таких как Docker Hub, для хранения и обмена образами контейнеров в экосистеме разработки машинного обучения и глубокого обучения?

Каковы недостатки использования контейнеров для развертывания приложений машинного обучения и глубокого обучения, такие как сложность управления, стоимость и ограничения безопасности?

Опишите процесс миграции существующего приложения машинного обучения или глубокого обучения на использование контейнеров для обеспечения масштабируемости, повторного использования кода и переносимости.

Сравните и сопоставьте подходы к управлению ресурсами, планированию и оркестрации для контейнеров, используемые в различных системах, таких как Kubernetes, Mesos и Docker Swarm.

Как контейнеры облегчают развертывание моделей машинного обучения и глубокого обучения в производство?  
 В чём заключается роль систем видеоаналитики в современных приложениях и инфраструктурах?  
 Дайте обзор фреймворка DeepStream и опишите его основные возможности.  
 Какие задачи компьютерного зрения могут быть решены с использованием DeepStream?  
 Сравните фреймворк DeepStream с другими решениями для видеоаналитики, такими как OpenVINO и Darknet.  
 Каким образом DeepStream интегрируется с платформами и системами видеонаблюдения, такими как Intel® OpenVINO™ и ONVIF?  
 Охарактеризуйте процесс обработки видео с использованием DeepStream, включая этапы обнаружения объектов, классификации и отслеживания.  
 Какова роль сверточных нейронных сетей (CNN) и глубокого обучения (DL) в алгоритмах видеоаналитики на базе DeepStream?  
 Как обеспечивается масштабируемость и гибкость систем видеоаналитики с использованием DeepStream в распределенных и облачных инфраструктурах?  
 Каковы преимущества использования DeepStream для развертывания систем видеоаналитики по сравнению с традиционными решениями?  
 Опишите процесс внедрения и интеграции DeepStream в существующие системы видеонаблюдения и машинного зрения.

#### 5.4. Методика оценки освоения дисциплины (модуля, практики. НР)

Оценка за дисциплину выставляется на основе результатов выполнения электронных тестов. Максимальное количество баллов получаемых за каждый тест равно 100.  
 Количество вопросов в тесте 25 по 4 балла за вопрос. Тесты сдаются только после выполнения практического задания и выполняются один раз.  
 Вся дисциплина оценивается в 1000 баллов. Оценкам соответствуют следующие баллы: «отлично» - 1000 - 901, «хорошо» - 900 - 750, «удовлетворительно» 749 – 600.

### 6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ

#### 6.1. Рекомендуемая литература

##### 6.1.1. Основная литература

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л1.1	Барский А. Б.	Логические нейронные сети: учебное пособие	Электронная библиотека	Москва: Интернет-Университет Информационных Технологий (ИНТУИТ) Бином. Лаборатория знаний, 2007
Л1.2	Петровичев Е. И.	Введение в искусственные нейронные сети: учеб. пособие по дисц. "Нейротехнологии в управлении"	Библиотека МИСиС	М.: Изд-во МГТУ, 2008
Л1.3	Галушкин А. И.	Нейронные сети: основы теории: монография	Библиотека МИСиС	М.: Горячая линия - Телеком, 2010
Л1.4	Барский А. Б.	Введение в нейронные сети: практическое пособие	Электронная библиотека	Москва: Интернет-Университет Информационных Технологий (ИНТУИТ),

#### 6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

Э1	Image Classification with DIGITS	<a href="https://courses.nvidia.com/courses/course-v1:DLI+L-FX-01+V1/about">https://courses.nvidia.com/courses/course-v1:DLI+L-FX-01+V1/about</a>
Э2	AWS for Deep Learning	<a href="https://aws.amazon.com/ru/console/">https://aws.amazon.com/ru/console/</a>
Э3	Площадка для выполнения практических заданий google	<a href="https://colab.research.google.com/">https://colab.research.google.com/</a>
Э4	Getting Started with Deep Learning	<a href="https://courses.nvidia.com/courses/course-v1:DLI+S-FX-01+V1/about">https://courses.nvidia.com/courses/course-v1:DLI+S-FX-01+V1/about</a>
Э5	Getting Started with Image Segmentation	<a href="https://courses.nvidia.com/courses/course-v1:DLI+L-FX-04+V2/about">https://courses.nvidia.com/courses/course-v1:DLI+L-FX-04+V2/about</a>
Э6	Deep Learning at Scale with Horovod	<a href="https://courses.nvidia.com/courses/course-v1:DLI+L-FX-23+V2/about">https://courses.nvidia.com/courses/course-v1:DLI+L-FX-23+V2/about</a>
Э7	Optimization and Deployment of TensorFlow Models with TensorRT	<a href="https://courses.nvidia.com/courses/course-v1:DLI+L-FX-18+V2/about">https://courses.nvidia.com/courses/course-v1:DLI+L-FX-18+V2/about</a>
Э8	High-Performance Computing with Containers	<a href="https://courses.nvidia.com/courses/course-v1:DLI+L-AC-25+V1/about">https://courses.nvidia.com/courses/course-v1:DLI+L-AC-25+V1/about</a>

Э9	AI Workflows for Intelligent Video Analytics with DeepStream	<a href="https://courses.nvidia.com/courses/course-v1:DLI+L-IV-04+V1/about">https://courses.nvidia.com/courses/course-v1:DLI+L-IV-04+V1/about</a>
<b>6.3 Перечень программного обеспечения</b>		
П.1	Anaconda	
П.2	Oracle VM VirtualBox	
П.3	Python	
П.4	MS Teams	
<b>6.4. Перечень информационных справочных систем и профессиональных баз данных</b>		
И.1	<a href="https://courses.nvidia.com/">https://courses.nvidia.com/</a>	
И.2	<a href="https://aws.amazon.com/ru/console/">https://aws.amazon.com/ru/console/</a>	

## 7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ

Ауд.	Назначение	Оснащение
Б-902	Учебная аудитория	12 стационарных компьютеров (2 x core i5-3470 8gb RAM, 10 x ryzen5 2400g 32gb RAM) , пакет лицензионных программ MS Office, демонстрационное оборудование: доска , проектор мультимедийный, комплект учебной мебели на 19 мест
Б-907	Учебная аудитория	1 стационарный компьютер , пакет лицензионных программ MS Office, комплект учебной мебели на 42 посадочных места , демонстрационное оборудование: доска , проектор мультимедийный x 2 , экран x 2 , колонки
Читальный зал №3 (Б)	Аудитория для самостоятельной работы	комплект учебной мебели на 44 места для обучающихся, МФУ Xerox VersaLink B7025 с функцией масштабирования текстов и изображений, 8 ПК с доступом к ИТС «Интернет», ЭИОС университета через личный кабинет на платформе LMS Canvas, лицензионные программы MS Office, MS Teams, ESET Antivirus.

## 8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ

1. Лекции читаются в аудиториях и, одновременно, дистанционно в MS Teams.
2. Практические занятия проводятся в аудиториях кафедры инженерной кибернетики института ИТКН.
3. Для самостоятельной работы используется электронная обучающая система Canvas.
4. Консультации по курсу проводятся преподавателем по календарному плану графику в аудиториях кафедры, а также с использованием средств дистанционного общения электронной обучающей системы Canvas.
5. Текущий контроль проводится как в электронной форме на компьютерах в дисплейных классах кафедры с использованием электронной обучающей системы Canvas, а также электронных тестов, и в очной форме на занятиях в аудиториях кафедры. Для подготовки к контрольным мероприятиям студенту выдается перечень тем, по материалу которых будет контрольное мероприятие. Подготовить к контрольным тестам студента возможно при консультациях в системе смешанного обучения Canvas и консультациях на очных занятиях.
6. Обучение организуется в соответствии с настоящей программой. Самостоятельная работа студентов организуется и контролируется с помощью электронных версий конспекта лекций и пособий с вопросами для самопроверки, а также индивидуального опроса студентов во время проведения практических работ.
7. Для самостоятельной работы студентам предоставляются дисплейные классы библиотеки НИТУ МИСИС. В процессе самостоятельной работы студенты используют электронную обучающую систему Canvas.