

**Федеральное государственное образовательное бюджетное учреждение
высшего образования
«ФИНАНСОВЫЙ УНИВЕРСИТЕТ
ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»
(Финансовый университет)**

**Департамент анализа данных и машинного обучения
Факультета информационных технологий и анализа больших данных**

СОГЛАСОВАНО

Газпромбанк _____

(наименование организации)

Начальник управления алгоритмов

машинного обучения

(должность представителя работодателя)

_____ Дашковский В.А.

15.12.2022 г.

УТВЕРЖДАЮ

Проректор по учебной

и методической работе

_____ Е.А. Каменева

29.12.2022 г.

С.В. Макрушин, В.А. Малекова

Глубокое обучение

Рабочая программа дисциплины

для студентов, обучающихся по направлению подготовки

01.03.02 - Прикладная математика и информатика,

ОП «Прикладная математика и информатика»

(Анализ данных и принятие решений в экономике и финансах),

ОП «Прикладное машинное обучение» (Прикладное машинное обучение)

Рекомендовано Ученым советом

*Факультета информационных технологий и анализа больших данных
(протокол №27 от 15.12.2022 г.)*

Одобрено Советом учебно-научного

*Департамента анализа данных и машинного обучения
(протокол №6 от 13.12.2022 г.)*

Москва 2022

Оглавление

1.Наименование дисциплины.....	2
2.Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине	2
3.Место дисциплины в структуре образовательных программ.....	4
4.Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся	4
5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий.....	5
5.1.Содержание дисциплины.....	5
5.2.Учебно–тематический план.....	8
5.3. Содержание семинаров, практических занятий.....	10
6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине.....	111
6.1.Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы.....	11
6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю	133
7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине.....	14
8.Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины	20
9.Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины.....	20
10. Методические указания для обучающихся по освоению дисциплины.	21
11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем	23
12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине.....	23

1.Наименование дисциплины

«Глубокое обучение».

2.Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
ОП «Прикладная математика и информатика» (Анализ данных и принятие решений в экономике и финансах)			
ОПК-4	Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	1.Использует информационно-коммуникационные технологии при поиске необходимой информации, сбора, визуализации и обработки данных.	Знать: подходы глубокого обучения для обработки данных. Уметь: строить модели глубокого обучения для обработки данных.
		2.Осуществляет рациональный выбор программного продукта в зависимости от поставленной задачи.	Знать: различные модели глубокого обучения, их плюсы и минусы. Уметь: выбирать модели глубокого обучения, подходящие для решения имеющейся задачи.
		3.Владеет навыками обеспечения информационной безопасности автоматизированных систем.	Знать: риски использования моделей глубокого обучения при использовании в агентных средах с носителями естественного интеллекта. Уметь: предотвращать риски, вызванные эмоциональной рефлексией носителей естественного интеллекта из-за доминирования моделей глубокого обучения.
ПКП-5	Способность применять методы и инструменты анализа данных и машинного обучения при подготовке аналитического обоснования финансово-	1.Осуществляет обоснованный выбор методов и инструментов обработки данных для решения экономических и финансовых задач.	Знать: различные методы глубокого обучения. Уметь: обоснованно выбирать модели глубокого обучения, подходящие для решения имеющихся экономических и финансовых задач.
		2.Владеет навыками решения прикладных задач с использованием методов и инструментов анализа	Знать: прикладные методы создания моделей глубокого обучения. Уметь: создавать прикладные

	экономических решений	данных и машинного обучения.	модели глубокого обучения с помощью специализированных прикладных инструментов.
ОП «Прикладное машинное обучение» (Прикладное машинное обучение)			
ПКП-1	Способность собирать наборы данных, в том числе больших данных, выполнять их подготовку для анализа в соответствии с решаемой прикладной задачей	1. Владеет навыками поиска внешних и внутренних источников данных для решения прикладной задачи.	Знать: источники данных для обучения моделей глубокого обучения. Уметь: находить подходящие источники данных и адаптировать их для построенных прикладных моделей.
		2. Использует инструментальные средства для извлечения, преобразования, хранения и обработки данных из разнородных источников.	Знать: инструментальные средства для извлечения, преобразования, хранения и обработки данных из разнородных источников Уметь: использовать специализированные инструментальные средства для извлечения, преобразования, хранения и обработки данных для их применения в моделях глубокого обучения.
		3. Владеет навыками обработки больших данных для решения прикладных задач машинного обучения.	Знать: методы обработки больших данных для решения прикладных задач глубокого обучения. Уметь: использовать методы обработки больших данных для решения прикладных задач глубокого обучения.
ПКП-5	Способность строить, обучать и оценивать качество моделей глубокого обучения в прикладных задачах	1. Демонстрирует знание принципов построения искусственных нейронных сетей, их обучения и оценки качества.	Знать: принципы построения искусственных нейронных сетей. Уметь: обучать искусственные нейронные сети и оценивать их качество.
		2. Демонстрирует знание современных архитектур искусственных нейронных сетей.	Знать: современные архитектуры искусственных нейронных сетей. Уметь: обучать современные архитектуры искусственных нейронных сетей.
		3. Использует современные инструментальные средства для построения искусственных нейронных сетей для решения прикладных задач.	Знать: современные инструментальные средства для построения искусственных нейронных сетей. Уметь: использовать современные инструментальные средства для построения искусственных нейронных сетей.

3. Место дисциплины в структуре образовательных программ

Дисциплина «Глубокое обучение» является дисциплиной Профиля Анализ данных и принятие решений в экономике и финансах для направления подготовки 01.03.02 – Прикладная математика и информатика, ОП «Прикладная математика и информатика».

Дисциплина «Глубокое обучение» является дисциплиной Профиля Прикладное машинное обучение для направления подготовки 01.03.02 – Прикладная математика и информатика, ОП «Прикладное машинное обучение».

4. Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся

ОП «Прикладная математика и информатика» (Анализ данных и принятие решений в экономике и финансах)

Вид учебной работы по дисциплине	Всего (в з/е и часах)	Семестр 5 (в часах)
Общая трудоемкость дисциплины	4/144	144
Контактная работа – Аудиторные занятия	68	68
<i>Лекции</i>	16	16
<i>Семинары, практические занятия</i>	52	52
Самостоятельная работа	76	76
Вид текущего контроля	Контрольная работа	Контрольная работа
Вид промежуточной аттестации	Экзамен	Экзамен

ОП «Прикладное машинное обучение» (Прикладное машинное обучение)

Вид учебной работы по дисциплине	Всего (в з/е и часах)	Семестр 5 (в часах)
Общая трудоемкость дисциплины	5/180	180
Контактная работа – Аудиторные занятия	68	68
<i>Лекции</i>	16	16
<i>Семинары, практические занятия</i>	52	52
Самостоятельная работа	112	112
Вид текущего контроля	Контрольная работа	Контрольная работа
Вид промежуточной аттестации	Экзамен	Экзамен

5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий

5.1. Содержание дисциплины

Тема1. Введение в искусственные нейронные сети и глубокое обучение.

Знакомство с историей изучения искусственных нейронных сетей (ИНС) и глубокого обучения. Первая и вторая зима ИИ. Вторая весна ИИ и глубокое обучение. Предпосылки для наступления второй весны ИИ.

Персепетрон. Схема работы, функция активации и ее варианты, проблема по Знакомство с базовыми понятиями ИНС.

Принципы построения моделей глубокого обучения.

Основы построения моделей ИНС в PyTorch. Загрузка и преобразование данных в PyTorch. Тензоры и операции с ними в PyTorch: создание тензоров, операции с тензорами. Арифметические операции и математические функции: операции, изменяющие размер тензора, операции агрегации, матричные операции.

Тема 2. Обучение искусственных нейронных сетей.

Современные методы обучения нейронной сети и обратное распространение ошибки. Обучение модели нейронной сети, алгоритм обратного распространения ошибки. Проблема обучения модели нейронной сети. Проблема поиска градиента. Дифференцируемое программирование и реализация обратного распространения ошибки.

Современные техники глубокого обучения: улучшенные подходы к регуляризации, улучшенные схемы инициализации весов, улучшенные функции активации, улучшенные схемы оптимизации.

Автоматическое дифференцирование в PyTorch. Знакомство с основным циклом обучения в PyTorch на примере решения задачи регрессии.

Тема 3. Сверточные искусственные нейронные сети.

Представления о принципах обработки изображений в мозге человека. Специфика изображений как исходных данных.

Идея сверточных сетей и базовое описание работы одного сверточного слоя. Изменение размеров карты признаков: эффект границ и max-pooling. Применение одного слоя свертки к одному слою входных данных.

Иерархии сверточных сетей и типичная архитектура сети распознавания изображений.

Пример реализации архитектуры сверточной сети распознавания изображений на PyTorch.

Приемы для обучения глубокому обучению на небольших наборах изображений: аугментация, дообучение

Визуализация знаний, заключенных в сверточной нейронной сети

Тема 4. Представление объектов в виде векторов малой размерности.

Мотивация для построения эмбединга (представления объекта в виде векторов малой размерности). Метрики для векторных представлений объектов.

Векторное представление слов. Дистрибутивная гипотеза. Матрица совместной встречаемости. Представление слов в виде векторов малой размерности. Сингулярное разложение (Singular Value Decomposition, SVD). Алгоритм word2vec. Модели CBOW и Skip-Gram.

Базовая реализация CBOW на PyTorch.

Тема 5. Рекуррентные нейронные сети.

Базовые подходы к RNN (рекуррентным нейронным сетям): Мотивация для использования рекуррентных нейронных сетей, принцип работы рекуррентной нейронной сети, проблемы при обучении и применении рекуррентных нейронных сетей. Обратное распространение ошибки в архитектуре RNN.

Построение многослойных нейронных сетей на базе архитектуры RNN. Двухнаправленные рекуррентные сети (bidirectional RNN).

Архитектура ИНС LSTM (Long Short-Term Memory) и GRU (gated recurrent unit).

Пример реализации RNN на PyTorch. Пример модели ИНС на основе LSTM, реализованной на PyTorch.

Тема 6. Механизм внимания для искусственных нейронных сетей.

Введение в механизм внимания в ИНС.

Базовые принципы применения механизма внимания для задачи Seq2seq. Использование механизмов внимания для других задач.

Архитектура Transformer. Принцип работы автоэнкодера. Модуль Multi-head attention.

Языковая модель Bidirectional Encoder Representations from Transformers (BERT).

Пример использования BERT.

Тема 7. Обучение с подкреплением.

Постановка задачи обучения с подкреплением.

Подход на основе Policy Gradient. Подход на основе Q-learning. Подходы RL основанные на моделях.

Тема 8. Графовые нейронные сети

Введение в графовые нейронные сети. Проблематика графов как объектов для эмбединга. Общая схема построения графовых нейронных сетей без учителя. Задача semi-supervised learning.

Фреймворк Message Passing Neural Networks (MPNN). Фрэймворк GCN (сверточные нейронные сети для решеток и графов). Спектральные модели GNN.

Общая схема конвейера для построения модели GNN (графовой нейронной сети). Типология моделей распространения GNN.

5.2. Учебно-тематический план

ОП «Прикладная математика и информатика» (Анализ данных и принятие решений в экономике и финансах)

№ п/п	Наименование тем (разделов) дисциплины	Трудоемкость в часах					Формы текущего контроля успеваемости
		Всего	Контактная работа - Аудиторная работа			Самостоя тельная работа	
			Общая, в т.ч.:	Лекции	Семинары, практическ ие занятия		
1.	Введение в искусственные нейронные сети и глубокое обучение.	16	8	2	6	8	Самостоятельн ые работы. Участие в решении задач на практических занятиях. Собеседования по домашним заданиям.
2.	Обучение искусственных нейронных сетей.	18	8	2	6	10	
3.	Сверточные искусственные нейронные сети.	20	10	2	8	10	
4.	Представление объектов в виде векторов малой размерности.	20	10	2	8	10	
5.	Рекуррентные нейронные сети.	18	8	2	6	10	Самостоятельн ые работы. Участие в решении задач на практических занятиях. Собеседования по домашним заданиям.
6.	Механизм внимания для искусственных нейронных сетей.	18	8	2	6	10	
7.	Обучение с подкреплением.	16	8	2	6	8	
8.	Графовые нейронные сети.	18	8	2	6	10	
	В целом по дисциплине	144	68	16	52	76	Согласно учебному плану: контрольная работа
	Итого в %		47	24	76	53	

**ОП «Прикладное машинное обучение»
(Прикладное машинное обучение)**

№ п/п	Наименование тем (разделов) дисциплины	Трудоемкость в часах					Формы текущего контроля успеваемости
		Всего	Контактная работа - Аудиторная работа			Самостоят ельная работа	
			Общая, в т.ч.:	Лекции	Семинары, практическ ие занятия		
1.	Введение в искусственные нейронные сети и глубокое обучение.	22	8	2	6	14	Самостоятельн ые работы. Участие в решении задач на практических занятиях. Собеседования по домашним заданиям.
2.	Обучение искусственных нейронных сетей.	22	8	2	6	14	
3.	Сверточные искусственные нейронные сети.	24	10	2	8	14	
4.	Представление объектов в виде векторов малой размерности.	24	10	2	8	14	Самостоятельн ые работы. Участие в решении задач на практических занятиях. Собеседования по домашним заданиям.
5.	Рекуррентные нейронные сети.	22	8	2	6	14	
6.	Механизм внимания для искусственных нейронных сетей.	22	8	2	6	14	
7.	Обучение с подкреплением.	22	8	2	6	14	
8.	Графовые нейронные сети.	22	8	2	6	14	
	В целом по дисциплине	180	68	16	52	112	Согласно учебному плану: контрольная работа
	Итого в %		38	24	76	62	

5.3. Содержание семинаров, практических занятий

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарских, практических занятиях, рекомендуемые источники из разделов 8,9 (указывается раздел и порядковый номер источника)	Формы проведения занятий
Введение в искусственные нейронные сети и глубокое обучение.	Решение задач на освоение основе PyTorch и построения моделей ИНС в PyTorch. <i>Рекомендуемые источники: п.8, [1]- [4]</i>	Интерактивная форма, Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений
Обучение искусственных нейронных сетей.	Знакомство с основным циклом обучения в PyTorch на примере решения задачи регрессии. Решение других задач с использованием основного цикла обучения в PyTorch. <i>Рекомендуемые источники: п.8, [1]-[4]</i>	Интерактивная форма, Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений
Сверточные искусственные нейронные сети.	Решение задач классификации изображений с помощью архитектуры на основе сверточных ИНС на PyTorch. <i>Рекомендуемые источники: п.8, [1]-[4]</i>	Интерактивная форма, Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений
Представление объектов в виде векторов малой размерности.	Решение задач построения эмбедингов с помощью моделей word2vec. <i>Рекомендуемые источники: п.8, [1]-[4]</i>	Интерактивная форма, Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений
Рекуррентные нейронные сети.	Решение задач анализа последовательностей с помощью ИНС с архитектур RNN и LSTM. <i>Рекомендуемые источники: п.8, [1]-[4]</i>	Интерактивная форма, Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений

Механизм внимания для искусственных нейронных сетей.	Решение задач повышения качества моделей с помощью включения механизма внимания. Решение задач анализа текста на естественном языке с помощью BERT. <i>Рекомендуемые источники: п.8, [1]-[4]</i>	Интерактивная форма, Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений
Обучение с подкреплением.	Решение задач обучения с подкреплением с помощью готовых архитектур ИНС. <i>Рекомендуемые источники: п.8, [1]- [4]</i>	Интерактивная форма, Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений
Графовые нейронные сети.	Построение эмбединга узлов сети с помощью графовых нейронных сетей. <i>Рекомендуемые источники: п.8, [1]-[4]</i>	Интерактивная форма, Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений

6. Учебно-методическое обеспечение для самостоятельной работы обучающихся по дисциплине

6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы

Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
Введение в искусственные нейронные сети и глубокое обучение.	Решение задач на освоение основ PyTorch: построение различных моделей персептрона с помощью PyTorch.	Работа с учебной литературой. Решение типовых задач. Разбор вопросов по теме занятия. Выполнение домашних заданий к каждому занятию.

Обучение искусственных нейронных сетей.	Решение продвинутых задач с использованием основного цикла обучения в PyTorch, изучение современных техник глубокого обучения: улучшенных подходов к регуляризации, улучшенные схемы инициализации весов, улучшенные функции активации, улучшенные схемы оптимизации.	Работа с учебной литературой. Решение типовых задач. Разбор вопросов по теме занятия. Выполнение домашних заданий к каждому занятию.
Сверточные искусственные нейронные сети.	Решение задач классификации изображений с помощью архитектуры на основе сверточных ИНС на PyTorch: глубокое обучение на небольших наборах изображений с помощью аугментации и дообучения.	Работа с учебной литературой. Решение типовых задач. Разбор вопросов по теме занятия. Выполнение домашних заданий к каждому занятию.
Представление объектов в виде векторов малой размерности.	Решение задач построения эмбедингов с помощью моделей word2vec: обучение и дообучение моделей word2vec.	Работа с учебной литературой. Решение типовых задач. Разбор вопросов по теме занятия. Выполнение домашних заданий к каждому занятию.
Рекуррентные нейронные сети.	Решение задач анализа последовательностей с помощью ИНС с архитектуры GRU.	Работа с учебной литературой. Решение типовых задач. Разбор вопросов по теме занятия. Выполнение домашних заданий к каждому занятию.
Механизм внимания для искусственных нейронных сетей.	Решение задачи дообучения модели BERT.	Работа с учебной литературой. Решение типовых задач. Разбор вопросов по теме занятия. Выполнение домашних заданий к каждому занятию.
Обучение с подкреплением.	Доработка готовых архитектур для обучения с подкреплением.	Работа с учебной литературой. Решение типовых задач. Разбор вопросов по теме занятия. Выполнение домашних заданий к каждому занятию.
Графовые нейронные сети.	Построение эмбединга узлов сети с дополнительными признаками с помощью графовых нейронных сетей.	Работа с учебной литературой. Решение типовых задач. Разбор вопросов по теме занятия. Выполнение домашних заданий к каждому занятию.

6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю

Примерные вопросы к контрольной работе

1. Персепетрон. Схема работы, функция активации и ее варианты.
2. Загрузка и преобразование данных в PyTorch.
3. Тензоры и операции с ними в PyTorch: создание тензоров, операции с тензорами.
4. Современные методы обучения нейронной сети и обратное распространение ошибки.
5. Современные техники глубокого обучения: улучшенные подходы к регуляризации.
6. Современные техники глубокого обучения: улучшенные схемы инициализации весов.
7. Современные техники глубокого обучения: улучшенные функции активации.
8. Современные техники глубокого обучения: улучшенные схемы оптимизации.
9. Автоматическое дифференцирование в PyTorch.
10. Основной цикл обучения в PyTorch.
11. Сверточные сети и базовое описание работы одного сверточного слоя.
12. Иерархия сверточных сетей и типичная архитектура сети распознавания изображений.

Примерные задания контрольной работы

1. Какие виды активационных функций применяются в искусственных нейронных сетях?
2. Опишите механизм обратного распространения ошибки и принцип обучения глубоких ИНС с его помощью.
3. Минбатчи – что из себя представляют и какова причина их использования.
4. Проблема инициализации весов при обучении ИНС. Инициализация Ксавье.
5. Опишите архитектуру модуля LSTM
6. Опишите архитектуру модели CBOW в Word2vec и принципы ее обучения и применения.

7. Опишите принцип работы слоя регуляризации Dropout и причины его использования.
8. Опишите преобразование Softmax и ситуации для его применения в глубоком обучении.
9. Опишите функцию потерь Cross Entropy loss и принцип ее использования.
10. Модель перцептрона. Проблема линейно неразделимых множеств и ее решение.

Критерии балльной оценки различных форм текущего контроля успеваемости содержатся в соответствующих методических рекомендациях Департамента анализа данных и машинного обучения Факультета информационных технологий и анализа больших данных.

7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине

Перечень компетенций с указанием индикаторов их достижения в процессе освоения образовательной программы содержится в разделе 2. «Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине».

Типовые контрольные задания или иные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний

Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания

ОП «Прикладная математика и информатика» (Анализ данных и принятие решений в экономике и финансах)			
ОПК-4 Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности	1.Использует информационно-коммуникационные технологии при поиске необходимой информации, сбора, визуализации и обработки данных.	Знать: подходы глубокого обучения для обработки данных. Уметь: строить модели глубокого обучения для обработки данных.	Постройте модель глубокого обучения для классификации рукописных цифр.
	2.Осуществляет рациональный выбор программного продукта в зависимости от поставленной задачи.	Знать: различные модели глубокого обучения, их плюсы и минусы. Уметь: выбирать модели глубокого обучения, подходящие для решения имеющейся задачи.	Определите архитектуру ИНС для задачи классификации рукописных цифр.
	3. Владеет навыками обеспечения информационной безопасности автоматизированных систем.	Знать: риски использования моделей глубокого обучения при использовании в агентных средах с носителями естественного интеллекта. Уметь: предотвращать риски, вызванные эмоциональной рефлексией носителей естественного интеллекта из-за доминирования моделей глубокого обучения.	Укажите в каких областях модели ИНС превосходят человека по качеству решения задач с формальной метрикой.
ПКП-5 Способность применять методы и инструменты анализа данных и машинного обучения при подготовке аналитического обоснования финансово-экономических решений	1.Осуществляет обоснованный выбор методов и инструментов обработки данных для решения экономических и финансовых задач.	Знать: различные методы глубокого обучения. Уметь: обоснованно выбирать модели глубокого обучения, подходящие для решения имеющихся экономических и финансовых задач.	Определите архитектуру ИНС для задачи предсказания курса валюты на основе высокодискретного ряда для задачи краткосрочного прогнозирования..
	2.Владеет навыками решения прикладных задач с использованием методов и инструментов анализа данных и машинного обучения.	Знать: прикладные методы создания моделей глубокого обучения. Уметь: создавать прикладные модели глубокого обучения с помощью специализированных прикладных инструментов.	Реализуйте на PyTorch модель определения тональности короткого текста (до 256 символов).

ОП «Прикладное машинное обучение» (Прикладное машинное обучение)			
ПКП-1 Способность собирать наборы данных, в том числе больших данных, выполнять их подготовку для анализа в соответствии с решаемой прикладной задачей	1. Владеет навыками поиска внешних и внутренних источников данных для решения прикладной задачи.	Знать: источники данных для обучения моделей глубокого обучения. Уметь: находить подходящие источники данных и адаптировать их для построенных прикладных моделей.	Найдите и подготовьте датасет для обучения глубокой сверточной модели классификации изображений по большому количеству классов.
	2. Использует инструментальные средства для извлечения, преобразования, хранения и обработки данных из разнородных источников.	Знать: инструментальные средства для извлечения, преобразования, хранения и обработки данных из разнородных источников Уметь: использовать специализированные инструментальные средства для извлечения, преобразования, хранения и обработки данных для их применения в моделях глубокого обучения.	Преобразуйте большое количество текстов на естественном языке в датасет для обучения модели word2vec.
	3. Владеет навыками обработки больших данных для решения прикладных задач машинного обучения.	Знать: методы обработки больших данных для решения прикладных задач глубокого обучения. Уметь: использовать методы обработки больших данных для решения прикладных задач глубокого обучения.	Определите архитектуру решения для обучения модели глубокой ИНС на большом наборе данных с помощью параллельной архитектуры обработки данных.
ПКП-5 Способность строить, обучать и оценивать качество моделей глубокого обучения в прикладных задачах	1. Демонстрирует знание принципов построения искусственных нейронных сетей, их обучения и оценки качества.	Знать: принципы построения искусственных нейронных сетей. Уметь: обучать искусственные нейронные сети и оценивать их качество.	Постройте модель глубокого обучения для классификации рукописных цифр.
	2. Демонстрирует знание современных архитектур искусственных нейронных сетей.	Знать: современные архитектуры искусственных нейронных сетей. Уметь: обучать современные архитектуры искусственных нейронных сетей.	Предложите архитектуру для классификации изображения с рукописной цифрой.

	3.Использует современные инструментальные средства для построения искусственных нейронных сетей для решения прикладных задач.	<u>Знать:</u> современные инструментальные средства для построения искусственных нейронных сетей. <u>Уметь:</u> использовать современные инструментальные средства для построения искусственных нейронных сетей.	Реализуйте на PyTorch модель определения тональности короткого текста (до 256 символов).
--	---	---	--

Примерные вопросы для подготовки к экзамену

1. Модель перцептрона. Проблема линейно неразделимых множеств и ее решение. Логика построения многослойных ИНС.
2. Функции активации. Требования к функциям активации Популярные функции активации.
3. Глубокое обучение. «Вторая весна искусственного интеллекта» и ее причины.
4. Линейное отображение. Векторно-матричное дифференцирование.
5. Проблема поиска градиента в общей логике обучения нейронной сети. Градиент функции многих переменных. Методы вычисления.
6. Кросс-валидация. Выборки train, validation, test. Проблема переобучения. Ранняя остановка.
7. Преобразование Softmax и функция потерь Cross Entropy loss.
8. Механизм обратного распространения ошибки.
9. Дифференцируемое программирование и реализация обратного распространения ошибки.
10. Стохастический градиентный спуск. Батчи обучающей выборки.
11. Адаптивные методы градиентного спуска. Метод импульсов. Метод Нестерова.
12. Проблема инициализации весов при обучении ИНС. Инициализация Ксавье.
13. Гиперпараметры. Скорость обучения и размер батча.
14. Переобучение модели и регуляризация. Dropout.
15. Минбатчи – причина использования. Нормализация по мини-батчам.

16. Многослойные сети. Граф потока вычислений.
17. Специфика задач машинного обучения на изображениях. Принцип работы сверточных сетей. Преимущества сверточных сетей при решении этих задач.
18. Архитектура многослойной ИНС распознавания изображений на основе сверточных сетей.
19. Приемы для глубокого обучения на небольших наборах изображений.
20. Схема работы сверточной сети. Операции свертки, пулинга, общий вид сверточной сети для решения задачи классификации изображения.
21. Задачи обработки текста: дистрибутивная семантика, матрица совместной встречаемости, представление слов в виде векторов малой размерности.
22. Word2vec: модель CBOW.
23. Word2vec: модель Skip-Gram.
24. Рекуррентная нейронная сеть, принципы ее обучения. Сложности применения рекуррентных нейронных сетей.
25. Модуль LSTM.
26. Механизм Attention. Пример использования Attention.
27. Архитектура Transformer.
28. Модель BERT.
29. Класс Tensor. Операции, изменяющие размер тензора. Операции агрегации.
30. Принципиальная логика обучения нейронной сети.
31. Автоматическое дифференцирование в PyTorch. Пример и применение в обучении ИНС.
32. Загрузка и преобразование данных. Классы Dataset, DataLoader, Transforms (и композиция трансформеров).
33. Класс nn.Module. Назначение. Основные поля, методы.
34. Линейные слои (Linear Layers).
35. Слои нелинейной активации (Non Linear Activations).
36. Слои нормализации (Normalization Layers).
37. Слои регуляризации (Dropout Layers).
38. Сверточные слои (Convolution Layers). Сжимающие слои (Pooling Layers).

- 39.Слои функций потерь (Loss Functions).
- 40.Слои эмбедингов nn.Embedding и их применение.
- 41.Класс torch.nn.LSTM и torch.nn.GRU.

Пример экзаменационного билета

**Федеральное государственное образовательное бюджетное учреждение
высшего образования**

**«ФИНАНСОВЫЙ УНИВЕРСИТЕТ ПРИ ПРАВИТЕЛЬСТВЕ
РОССИЙСКОЙ ФЕДЕРАЦИИ»
(Финансовый университет)**

Департамент анализа данных и машинного обучения

Дисциплина **Глубокое обучение**

Факультет информационных технологий и анализа больших данных

Форма обучения: **очная**

Направление подготовки: **01.03.02 Прикладная математика и информатика**

Направленность программы: _____

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №

1. Принципиальная логика обучения нейронной сети. (20 баллов)
2. На вход ИНС подаётся картинка 3x100x100. К нему применяются подряд 3 свёртки 3x3 с 64 фильтрами, max-pooling 3x3, еще 3 свёртки 3x3 с 128 нейронами. Определить какой будет размер тензора на каждом преобразовании при батче 64. Какой размер будет у последнего слоя? Какой слой труднее всего вычислять? Padding не используется, stride тоже. (20 баллов)
3. Реализовав сверточную нейронную сеть при помощи библиотеки PyTorch, решите задачу классификации изображений. Разделите набор данных на обучающее и тестовое множество. Выполните предобработку данных (приведите изображения к одному размеру и преобразуйте в тензоры). Отобразите графики значений функции потерь на обучающем множестве. Отобразите confusion matrix и classification report, рассчитанные на основе тестового множества. Датасет: images/eng_handwritten. (20 баллов)

8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины:

Основная литература:

- 1.Криволапов, С. Я. Математика на Python : учебник / С. Я. Криволапов, М. Б. Хрипунова. — Москва : КноРус, 2022. — 455. — ЭБС BOOK.ru. - URL:<https://book.ru/book/943665> (дата обращения: 10.03.2023). — Текст : электронный.
- 2.Криволапов, С. Я. Использование языка Python в теории вероятностей : учебник / С. Я. Криволапов. — Москва : Прометей, 2022. — 492 с. — ЭБС Лань. — URL: <https://e.lanbook.com/book/220814> (дата обращения: 10.03.2023). — Текст: электронный.

Дополнительная литература:

- 3.Криволапов, С. Я. Статистические вычисления на платформе Jupyter Notebook с использованием Python: учебник / С. Я. Криволапов. — Москва: КноРус, 2022. — 431 с. — ЭБС BOOK.RU. - URL:<https://book.ru/book/943660> (дата обращения: 10.03.2023). — Текст : электронный.
- 4.Ростовцев, В. С. Искусственные нейронные сети : учебник / В. С. Ростовцев. — Санкт-Петербург : Лань, 2021. — 216 с. — ЭБС Лань. — URL: <https://e.lanbook.com/reader/book/160142> (дата обращения: 10.03.2023). — Текст: электронный.

9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины:

1. Электронная библиотека Финансового университета (ЭБ) <http://elb.fa.ru/>
2. Электронно-библиотечная система BOOK.RU <http://www.book.ru>
3. Электронно-библиотечная система «Университетская библиотека ОНЛАЙН» <http://bibliodub.ru/>
4. Электронно-библиотечная система Znani um <http://www.znani um.com>
- 5.Электронно-библиотечная система издательства «ЮРАЙТ» <https://urait.ru/>
- 6.Электронно-библиотечная система издательства Проспект

<http://ebs.prospekt.org/books>

7. Электронно-библиотечная система издательства Лань <https://e.lanbook.com/>

8. Деловая онлайн-библиотека Alpina Digital <http://lib.alpinadigital.ru/>

9. Электронная библиотека Издательского дома «Гребенников» <https://grebennikon.ru/>

8. Научная электронная библиотека eLibrary.ru <http://elibrary.ru>

10. Национальная электронная библиотека <http://нэб.рф/>

11. Финансовая справочная система «Финансовый директор» <http://www.1fd.ru/>

12. Электронная документация по работе в pytorch (вариант 1):

<https://neurohive.io/ru/tutorial/glubokoeobucheniye-s-pytorch/>

13. Электронная документация по работе в pytorch (вариант 2):

<https://pytorch.org/tutorials/>

14. Дистрибутив anaconda: <http://anaconda.org/>

15. Электронная документация по работе в библиотеке pandas:

<http://pandas.pydata.org/>

16. Курс лекций по машинному и глубокому обучению (электронный учебник):

<http://www.machinelearning.ru/wiki/images/6/6d/Voron-ML-1.pdf>

10. Методические указания для обучающихся по освоению дисциплины

Основные этапы работы студента по дисциплине *Глубокое обучение*:

1. Предварительная ориентировка в подлежащем изучению учебном материале по программе.
2. Ознакомление с рекомендованной учебной литературой.
3. Слушание и конспектирование лекций, а также выполнение других видов учебной работы.
4. Планирование самостоятельной работы.
5. Обобщение и систематизация информации, почерпнутой из лекций и прочитанной литературы.
6. Выполнение контрольной работы.

Рекомендации по работе с учебным материалом:

1. Осознавайте наличный уровень полученных вами знаний.
2. В ситуации непонимания нужно выявить тот первичный уровень и факторы непонимания, которые стали препятствием понимания последующего.
3. Задавайте сами себе вопросы и пытайтесь ответить на них.

Рекомендации по работе на лекции и с лекционным материалом:

1. Основная задача на лекции – осмысление излагаемого в ней материала. Для этого необходимо слушать лекцию с самого начала, не упуская общих, ориентирующих в материале рассуждений и установок лектора.
2. Ведение записей на лекции важно и полезно для лучшего осмысливания материала, для сохранения информации, с целью ее дальнейшего использования.
3. Для облегчения записи рекомендуется применять сокращения повторяющихся терминов или хорошо известных понятий.

Рекомендации по работе с литературой:

1. Если возникли затруднения при разыскивании материала, по какому-либо конкретному вопросу, следует обратиться к предметному указателю, напечатанному, как правило, в конце каждого литературного источника.
2. Предметный указатель – это алфавитный список основных научных понятий (терминов), содержание которых раскрыто в книге, рядом с термином стоят числа, обозначающие номера страниц, на которых изложен материал, относящийся к данному понятию.

Рекомендации по выполнению контрольной работы:

1. Перед выполнением контрольной работы студент должен изучить соответствующие разделы учебной литературы.
2. Контрольную работу студент должен выполнять самостоятельно, используя те навыки и умения, которые получил на лекциях и практических занятиях.
3. При затруднениях, возникших при выполнении контрольной работы, студент может получить консультацию преподавателя.

11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем

11.1 Комплект лицензионного программного обеспечения:

1. Пакет офисных программ
2. Антивирус Kaspersky

11.2 Современные профессиональные базы данных и информационные справочные системы

1. Информационно-правовая система «Гарант»,
2. Информационно-правовая система «Консультант Плюс»,
3. Электронная энциклопедия: <http://ru.wikipedia.org/wiki/Wiki>
4. Система комплексного раскрытия информации «СКРИН» - <http://www.skrin.ru/>

11.3. Сертифицированные программные и аппаратные средства защиты информации:

- не предусмотрены.

11.4. Язык программирования Python 3.x в среде Windows.

11.5. Платформа для научных исследований, основанная на языке программирования Python, Anaconda, PyTorch.

12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Наличие аудитории, оснащенной компьютерной техникой и проектором, с возможностью подключения к сети «Интернет».