

**Федеральное государственное образовательное бюджетное
учреждение высшего образования
«ФИНАНСОВЫЙ УНИВЕРСИТЕТ
ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»
(Финансовый университет)
Департамент бизнес-информатики
Факультета информационных технологий и анализа больших данных**

УТВЕРЖДАЮ

Проректор по учебной
и методической работе
_____ Е.А. Каменева
25.12.2023 г.

П.А. Сахнюк

Искусственный интеллект в государственном управлении

Рабочая программа дисциплины
для студентов, обучающихся по направлению подготовки
38.03.04 - Государственное и муниципальное управление
ОП «Цифровое государство и экономика»

*Рекомендовано Ученым советом Факультета информационных
технологий и анализа больших данных
(протокол №39 от 20 декабря 2023 г.)*

*Одобрено Советом учебно-научного Департамента бизнес-
информатики
(протокол № 3 от 18 декабря 2023 г.)*

Москва 2023

Содержание

1. Наименование дисциплины.....	3
2. Перечень планируемых результатов освоения образовательной программы с указанием индикаторов их достижения, соотнесенных с планируемыми результатами обучения по дисциплине.....	3
3. Место дисциплины в структуре образовательной программы.....	7
4. Объем дисциплины в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся.....	7
5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий.....	8
5.1. Содержание дисциплины.....	8
5.2. Учебно-тематический план.....	10
5.3. Содержание семинаров, практических занятий.....	10
6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине.....	12
6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы.....	12
6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю.....	13
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.....	21
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.....	21
12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине.....	22

1. Наименование дисциплины

«Искусственный интеллект в государственном управлении».

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

В результате изучения дисциплины у студентов должны быть сформированы следующие компетенции:

Таблица 1

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (владения, умения и знания), соотнесенные с компетенциями/индикаторами достижения компетенции
УК-4	Способность использовать прикладное программное обеспечение при решении профессиональных задач	Использует основные методы и средства получения, представления, хранения и обработки данных	Знать: Основные методы получения данных (способы сбора данных, методы обработки данных, виды данных), методы представления данных (структурированные и неструктурированные), методы хранения данных и обработки данных (алгоритмы машинного обучения, методы выбора моделей) Уметь: Получать данные из различных источников, очищать и преобразовывать их в формат, подходящий для машинного обучения, представлять данные в формате, который может быть легко обработан алгоритмами машинного обучения, хранить данные в базах данных, применять алгоритмы машинного обучения для обработки данных.
		Демонстрирует владение профессиональными пакетами прикладных программ.	Знать: Основные структуры данных Python, фреймворки PyTorch, TensorFlow, Keras, и другие фреймворки для машинного обучения в Python, библиотеки для обработки данных, такие как NumPy, Pandas, Scikit-learn и другие, библиотеки визуализации данных Matplotlib, Seaborn и Plotly. Уметь: Использовать Python для решения задач машинного обучения, включая предобработку данных, моделирование с использованием

			NumPy, Pandas, Scikit-learn и др. библиотек python, выбор оптимальных параметров и оценку моделей, использовать фреймворки для машинного обучения, такие как PyTorch, TensorFlow, Keras, AutoML для создания и тренировки моделей.
		Выбирает необходимое прикладное программное обеспечение в зависимости от решаемой задачи.	Знать: Принципы и методы машинного обучения и искусственного интеллекта, чтобы понять, какие задачи лучше решать с помощью этих технологий, спектр доступных пакетов прикладных программ для машинного обучения и искусственного интеллекта, таких как Python (NumPy, Pandas, Scikit-learn, Keras, TensorFlow, PyTorch), а также другие языки и фреймворки, преимущества и недостатки различных пакетов прикладных программ для машинного обучения и искусственного интеллекта в зависимости от конкретной задачи, факторы, которые влияют на выбор пакета прикладных программ для машинного обучения и искусственного интеллекта, такие как объём данных, структура данных, требования к точности и скорости вычислений. Уметь: Использовать Python (NumPy, Pandas, Scikit-learn, Keras, TensorFlow, PyTorch и др.) для решения конкретных задач машинного обучения и искусственного интеллекта, выбирать пакет прикладных программ в зависимости от конкретной задачи, учитывая объём данных, структуру данных и требования к точности и скорости вычислений.

		Использует прикладное программное обеспечение для решения конкретных прикладных задач.	Знать: Преимущества и недостатки различных пакетов прикладных программ для машинного обучения и искусственного интеллекта в зависимости от конкретной прикладной задачи, методы преобразования данных для использования в моделях машинного обучения и искусственного интеллекта, метрики качества для оценки результатов моделей машинного обучения и искусственного интеллекта, способы использования моделей машинного обучения и искусственного интеллекта в реальных проектах. Уметь: Выбрать подходящий пакет прикладных программ для машинного обучения и искусственного интеллекта в зависимости от конкретной задачи, учитывая объем данных, структуру данных и требования к точности и скорости вычислений, преобразовать данные для использования в моделях машинного обучения и искусственного интеллекта, проводить анализ данных для выявления тенденций и обработки излишней информации, использовать модели машинного обучения и искусственного интеллекта в реальных проектах.
УК-15	Способность релевантно решаемым задачам использовать информационные ресурсы и информационно-коммуникационные технологии для достижения целей, связанных с профессиональной деятельностью, обучением, участием в жизни	1. Самостоятельно выбирает и использует цифровые средства общения, осуществляет поиск и/или создание контента в соответствии с целью взаимодействия, в том числе для организации совместной деятельности.	Знать: Как использовать цифровые средства общения для организации совместной деятельности в области машинного обучения и искусственного интеллекта, ресурсы для повышения своих навыков в области машинного обучения и искусственного интеллекта, (видеокурсы, блоги, книги и конференции), способы создания контента для обмена знаниями в области машинного обучения и искусственного интеллекта. Уметь: Использовать цифровые средства общения для обмена данными и

	общества и других сферах жизни		коллаборации на проектах в области машинного обучения и искусственного интеллекта, находить и использовать ресурсы для повышения своих навыков в области машинного обучения и искусственного интеллекта, создавать контент для обмена знаниями в области машинного обучения и искусственного интеллекта.
		2. Владеет навыками организации взаимодействия и коммуникации с помощью информационных систем и/или цифровых сервисов и технологий.	Знать: Методы использования цифровых средств общения для обмена данными и коллаборации на проектах в области машинного обучения и искусственного интеллекта, как использовать цифровые средства общения для коммуникации с экспертами в области машинного обучения и искусственного интеллекта для получения консультаций и решений проблем. Уметь: Использовать цифровые средства общения для обмена данными и коллаборации на проектах в области машинного обучения и искусственного интеллекта, использовать цифровые средства общения для коммуникации с экспертами в области машинного обучения и искусственного интеллекта для получения консультаций и решений проблем.
		3. Осуществляет подбор и применение различных информационно-коммуникационных средств для решения образовательных и профессиональных задач.	Знать: Преимущества и недостатки различных информационно-коммуникационных средств для решения задач в области искусственного интеллекта в зависимости от конкретной прикладной задачи. Уметь: Использовать различные информационно-коммуникационные средства для решения задач в области искусственного интеллекта в зависимости от конкретной прикладной задачи.

3. Место дисциплины в структуре образовательной программы

Дисциплина «Искусственный интеллект в государственном управлении» относится к циклу математики и информатики обязательной части ОП «Цифровое государство и экономика» по направлению подготовки 38.03.04 - Государственное и муниципальное управление/

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

Таблица 2

Вид учебной работы по дисциплине	Всего (в з/ед. и часах)	Семестр 3 (в часах)	Семестр 4 (в часах)
Общая трудоемкость дисциплины	6 ЗЕ./216 ч.	98	98
Контактная работа - Аудиторные занятия	100	50	50
<i>Лекции</i>	32	16	16
<i>Семинары, практические занятия</i>	68	34	34
Самостоятельная работа	116	58	58
Вид текущего контроля	Контрольная работа	Контрольная работа	-
Вид промежуточной аттестации	Зачет/Экзамен	Зачет	Экзамен

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

5.1. Содержание дисциплины

Тема 1. Машинное обучение на python

Применение pandas для анализа и обработки данных. Библиотеки визуализации данных Matplotlib, Seaborn, Altair, Plotly Express. Exploratory data analysis (EDA) - анализ основных свойств данных, нахождение в них общих закономерностей, распределений и аномалий, построение начальных моделей, зачастую с использованием инструментов визуализации. Автоматизация исследовательского анализа данных с использованием pandas profiling, библиотеки mitosheet, bamboolib, dataprep.

Методика CRISM-DM, машинное обучение для решения задач Data Mining. Градиентный спуск в машинном обучении. Деревья решений. Бэггинг, бустинг, стекинг. Фреймворки машинного обучения: XGBoost, LightGBM, CatBoost, h2o.ai. Кластерный анализ, алгоритм k-means. Поиск ассоциативных правил. Анализ временных рядов. Автоматическое машинное обучение (AutoML): rycaret, autogluon, LAMA.

Машинное обучение и технологии на базе искусственного интеллекта (ИИ) в платформах бизнес-аналитики и науки о данных. Применение машинного обучения на python в государственном управлении.

Тема 2. Нейронные сети. Платформы и фреймворки глубокого обучения

Введение в искусственные нейронные сети: биологический нейрон, модель Мак-Каллока-Питтса как линейный классификатор, функции активации, задача исключаящего или, теорема Колмогорова, алгоритм обратного распространения ошибок. Классификация изображений. Свёрточные сети (CNN). Архитектуры нейросетей. Современные архитектуры свёрточных сетей. Предотвращение переобучения. Transfer learning. Детекция и сегментация изображений.

Введение в NLP, идея эмбедингов: w2v. Seq2seq модели. Механизмы внимания. Автопереводчики. архитектура Transformer: Encoder-decoder, появление attention. Generative pre-trained transformer.

Решение задач Data Mining с помощью нейросетевых методов. Основные платформы и фреймворки глубокого обучения: TensorFlow, Keras, Hugging Face, fast.ai. Автоматизация задач Data Science за счет AutoML, AutoKeras, Autogluon. Критерии выбора фреймворка для глубокого обучения. Решение задач машинного обучения на неструктурированных и структурированных данных. Применение фреймворков глубокого обучения в государственном управлении.

Тема 3. Большие языковые модели

Генераторно-состязательная сети, большие языковые модели (Large Language Models - LLM). Типы больших языковых моделей: модель языкового

представления (Language Representation Model), модели с нулевым выстрелом (Zero-Shot Models), мультимодальные модели (Multimodal Models), специфичные для предметной области модели (Domain-Specific Models). Примеры больших языковых моделей: PaLM (Pathways Language Model), XLNet, BERT (Bidirectional Encoder Representations from Transformers), GPT (Generative Pretrained Transformers), LLaMA (Large Language Model by Meta AI).

Применение LLM Cloud.ru (Sbercloud.ru) GigaChat, RuGPT, YandexGPT, Hugging Face, OpenAI и др. для решения задач обработки естественного языка (NLP), компьютерного зрения (CV), обработки аудиоданных. Мультимодальные большие языковые модели: перевод текста, категоризация информации, определение эмоциональной окраски текста, создание новых текстов и предоставления ответов на вопросы, создание изображений и видео по текстовому описанию, создание чат ботов. Применение языковых моделей и сервисов для решения задач в государственном управлении.

5.2. Учебно-тематический план

Таблица 3

№ п/п	Наименование темы (раздела) дисциплины	Трудоемкость в часах					Формы текущего контроля успеваемости
		Всего	Контактная работа-Аудиторная работа			Самосто- ятельная работа	
			Общая	Лекции	Практ. и семинарские занятия		
3 семестр							
1	Тема 1. Машинное обучение на python	108	50	16	34	58	Выполнение индивидуальных заданий
	Итого в 3 семестре:	108	50	16	34	58	Контрольная работа
4 семестр							
2	Тема 2. Нейронные сети. Платформы и фреймворки глубокого обучения	56	26	8	18	30	Выполнение индивидуальных заданий
3	Тема 3. Большие языковые модели	52	24	8	16	28	Выполнение индивидуальных заданий
	Итого в 4 семестре:	108	50	16	34	58	
	В целом по дисциплине	216	100	32	68	116	
Итого в %:			46	32	68	54	

*объем контактной работы в очно-заочной/заочной формах обучения и индивидуальных учебных планах определяется соответствующими учебными планами. Темы, реализуемые в виде контактной работы, определяются преподавателем самостоятельно, исходя из уровня их сложности.

5.3. Содержание семинаров, практических занятий

Таблица 4

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
Тема 1. Машинное обучение на python	1. Решение регрессионной задачи с помощью библиотеки scikit-learn. 2. Исследование алгоритмов построения ансамблей деревьев решений. 3. Сравнение промышленных фреймворков машинного обучения. 4. Использование кластерного анализа для решения задач Data Mining. 5. Сравнение фреймворков машинного обучения, используемых для анализа и прогнозирования временных рядов. 6. Основные фреймворки машинного обучения с функцией AutoML.	Выполнение и защита практических заданий

	7. Разработка рекомендательной системы. 8. Применение технологий машинного обучения и искусственного интеллекта в платформах BI и Data Science. Основная литература: 1,2 Дополнительная литература: 1,2,3	
Тема 2. Нейронные сети. Платформы и фреймворки глубокого обучения	1. Введение в глубокое обучение 2. Первая модель в Colab 3. Обучение классификатора цифр 4. Классификация моделей одежды в Keras 5. Классификация изображений с TensorFlow 6. Трансферное обучение и тонкая настройка 7. Классификация изображений с Keras: классификация изображений с нуля, простая сеть MNIST, классификация изображений посредством точной настройки с помощью EfficientNet, классификация изображений с помощью Vision Transformer 8. Классификация и регрессия текста с AutoKeras 9. Классификация и регрессия с изображений с AutoKeras 10. Решение задач NLP с использованием библиотеки AutoGluon 11. Решение задач Computer Vision с использованием библиотеки AutoGluon Основная литература: 1,2 Дополнительная литература: 1,2,3	Выполнение и защита практических заданий
Тема 4. Большие языковые модели	1. Модели трансформеры: обработка естественного языка, трансформеры, что они могут, как работают трансформеры, модели энкодеров, модели декодеров, последовательные модели, предвзятость и ограничения 2. Использование трансформеров: токенизаторы, обработка нескольких последовательностей 3. Точная настройка предварительно обученной модели: обработка данных, точная настройка модели с помощью Trainer API или Keras 4. Совместное использование моделей и токенизаторов Hugging Face Hub, использование предварительно обученных моделей, совместное использование предварительно обученных моделей. 5. Применение мультимодальных больших языковых моделей (ChatGPT, Claude.AI, LLaMA, Perplexity, GigaChat, YandexGPT) в госуправлении. Основная литература: 1,2 Дополнительная литература: 1,2,3	Выполнение и защита практических заданий

6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

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

Таблица 5

Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
Тема 1. Машинное обучение на python	Использование алгоритма Apriori для решения задачи анализа рыночной корзины. Особенности фреймворков AutoML: FLAML; TPOT; Auto-sklearn.	Изучение методических материалов по теме в электронном виде и рекомендуемых разделов основной и дополнительной литературы, интернет – источников. Подготовка к семинарам
Тема 2. Нейронные сети. Платформы и фреймворки глубокого обучения	Фреймворки глубокого обучения: PyTorch, Lightning, Caffe2, Chainer, Microsoft Cognitive Toolkit, MxNet, PaddlePaddle.	Изучение методических материалов по теме в электронном виде и рекомендуемых разделов основной и дополнительной литературы, интернет – источников. Подготовка к семинарам
Тема 3. Большие языковые модели	AI Services: Yandex AI, clord.ru -Salute Speech - распознавание и синтез речи	Изучение методических материалов по теме в электронном виде и рекомендуемых разделов основной и дополнительной литературы, интернет – источников. Подготовка к семинарам

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

Примерные темы для контрольной работы:

1. Выбрать индивидуальное задание - использовать датасеты с Kaggle <https://www.kaggle.com/datasets?fileType=csv>
2. В Colab выполнить консолидацию данных кейса, если кейс состоит из нескольких
3. Провести очистку, трансформацию данных (ETL), выполнить preprocessing данных (<https://www.sv-europe.com/crisp-dm-methodology/>), провести EDA с использованием любой библиотеки python (легко использовать библиотеки python автоматизации EDA Dataprep, Autoviz: пример использования https://colab.research.google.com/drive/1PligDvYLxRVYzcBkcJ_ZQDqF4cLBWzhf?usp=sharing).
4. Обучить и оценить по метрикам качества модели машинного обучения (использовать различные алгоритмы); сравнить полученные результаты.
5. Обогащать данные кейса: провести ABC-XYZ, RFM-анализ, дополнить данные результатами решения задачи классификации, регрессии, кластеризации (одна или все на выбор исходя из данных кейса).
6. Экспортировать подготовленный датасет в BigQuery (или в формате .csv в Google Drive).
7. Обучить и сравнивать модели машинного обучения с использованием библиотеки scikit-learn (или LightGBM, XGBoost, CatBoost, h2o-3, AutoGluon, PyCaret используя алгоритмы регрессии, градиентного бустинга, случайного леса или используя платформы Loginom, Knime, RapidMiner). Провести прогнозирование ответа лучшей модели на новых данных (сгенерировать самостоятельно используя исходные данные или получить “новые” данные путем исключения из исходного датасета <https://colab.research.google.com/drive/1HTQJx8waY9OfzVejhJ--odaLIcHbLXYY?usp=sharing>).

Критерии балльной оценки различных форм текущего контроля успеваемости содержатся в соответствующих методических рекомендациях Департамента бизнес-информатики.

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

Перечень компетенций представлен в разделе 2, который характеризует перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы.

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

Таблица 6

Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (владения, умения и знания), соотнесенные с компетенциями/индикаторами достижения компетенции	Типовые контрольные задания
УК-4	1. Использует основные методы и средства получения, представления, хранения и обработки данных	Знать: Основные методы получения данных (способы сбора данных, методы обработки данных, виды данных), методы представления данных (структурированные и неструктурированные), методы хранения данных и обработки данных (алгоритмы машинного обучения, методы выбора моделей) Уметь: Получать данные из различных источников, очищать и преобразовывать их в формат, подходящий для машинного обучения, представлять данные в формате, который может быть легко обработан алгоритмами машинного обучения, хранить данные в базах данных, применять алгоритмы машинного обучения для обработки данных.	Задание 1 В репозитории TensorFlow Hub (содержащий предварительно обученные нейронные сети для решения многих видов задач, в основном в области компьютерного зрения и обработки естественного языка) найти модули для обработки текста для различных языков (английский, японский, немецкий, испанский, русский), word2vec, обученный на Википедии, NNLM эмбединг, обучить модель для классификации отзывов о фильмах на положительных и отрицательные, на небольшом наборе данных. Задание 2 В репозитории TensorFlow Hub найти модули для обработки изображений, используя метод трансферного обучения переобучить модель используя фреймворк Keras.
	2. Демонстрирует владение профессиональным и пакетами прикладных программ.	Знать: Основные структуры данных Python, фреймворки PyTorch, TensorFlow, Keras, и другие фреймворки для машинного обучения в Python, библиотеки для обработки данных, такие как NumPy, Pandas, Scikit-learn и другие, библиотеки визуализации данных Matplotlib, Seaborn и Plotly express. Уметь: Использовать Python для решения задач машинного обучения, включая	Задание 1 В Colab выполнить консолидацию данных кейса (если кейс состоит из нескольких таблиц), провести очистку, трансформацию данных (ETL), выполнить preprocessing данных, провести исследовательский анализ данных с использованием библиотек python Numpy, Pandas, Scikit-learn. (легко использовать библиотеки python автоматизации EDA Dataprep, AutoViz: пример использования Задание 2

		предобработку данных, моделирование с использованием Numpy, Pandas, Scikit-learn и др. библиотек python, выбор оптимальных параметров и оценку моделей, использовать фреймворки для машинного обучения, такие как PyTorch, TensorFlow, Keras, AutoML для создания и тренировки моделей.	В Colab провести очистку, трансформацию данных, выполнить preprocessing данных, провести исследовательский анализ данных с использованием библиотек визуализации данных Matplotlib, Seaborn и Plotly express
3. Выбирает необходимое прикладное программное обеспечение в зависимости от решаемой задачи.	Знать: Принципы и методы машинного обучения и искусственного интеллекта, чтобы понять, какие задачи лучше решать с помощью этих технологий, спектр доступных пакетов прикладных программ для машинного обучения и искусственного интеллекта, таких как Python (Numpy, Pandas, Scikit-learn, Keras, TensorFlow, PyTorch), а также другие языки и фреймворки, преимущества и недостатки различных пакетов прикладных программ для машинного обучения и искусственного интеллекта в зависимости от конкретной задачи, факторы, которые влияют на выбор пакета прикладных программ для машинного обучения и искусственного интеллекта, такие как объём данных, структура данных, требования к точности и скорости вычислений. Уметь: Использовать Python (NumPy, Pandas, Scikit-learn, Keras, TensorFlow, PyTorch и др.) для решения конкретных задач машинного обучения и искусственного интеллекта, выбирать пакет прикладных программ в зависимости от конкретной задачи, учитывая объём данных, структуру данных и требования к точности и скорости вычислений.	Задание 1 В Colab обучить и сравнивать модели машинного обучения с использованием фреймворков scikit-learn, LightGBM, XGBoost, CatBoost используя алгоритмы регрессии, градиентного бустинга, случайного леса или используя платформы. Провести прогнозирование ответа лучшей модели на новых данных Задание 2 В Colab обучить и сравнивать модели машинного обучения с использованием технологий AutoML фреймворков h2o-3, AutoGluon, PyCaret используя алгоритмы регрессии, градиентного бустинга, случайного леса или используя платформы. Провести прогнозирование ответа лучшей модели на новых данных	
4. Использует прикладное программное обеспечение для решения конкретных прикладных задач.	Знать: Преимущества и недостатки различных пакетов прикладных программ для машинного обучения и искусственного интеллекта в зависимости от конкретной задачи, факторы, которые влияют на выбор пакета прикладных программ для машинного обучения и искусственного интеллекта, такие как объём данных, структура данных, требования к точности и скорости вычислений.	Задание 1 В блокноте Jupyter (Jupyter Lab, Colab, Binder) на выбранном датасете провести разведочный анализ данных с использованием библиотеки Dataprep, D-Tale,	

		прикладной задачи, методы преобразования данных для использования в моделях машинного обучения и искусственного интеллекта, метрики качества для оценки результатов моделей машинного обучения и искусственного интеллекта, способы использования моделей машинного обучения и искусственного интеллекта в реальных проектах. Уметь: Выбрать подходящий пакет прикладных программ для машинного обучения и искусственного интеллекта в зависимости от конкретной задачи, учитывая объём данных, структуру данных и требования к точности и скорости вычислений, преобразовать данные для использования в моделях машинного обучения и искусственного интеллекта, проводить анализ данных для выявления тенденций и обработки излишней информации, использовать модели машинного обучения и искусственного интеллекта в реальных проектах.	Pandas-profiling. Используя фреймворки машинного обучения XGBoost, PyCaret создать, обучить и провести оценку модели машинного обучения. Задание 2 В блокноте Jupyter (Jupyter Lab, Colab, Binder) на выбранном датасете провести разведочный анализ данных с использованием библиотеки AutoViz, Sweetviz, dabl. Используя фреймворки машинного обучения CatBoost, LightGBM создать, обучить и провести оценку модели машинного обучения.
УК-15	1. Самостоятельно выбирает и использует цифровые средства общения, осуществляет поиск и/или создание контента в соответствии с целью взаимодействия, в том числе для организации совместной деятельности.	Знать: Как использовать цифровые средства общения для организации совместной деятельности в области машинного обучения и искусственного интеллекта, ресурсы для повышения своих навыков в области машинного обучения и искусственного интеллекта, (видеокурсы, блоги, книги и конференции), способы создания контента для обмена знаниями в области машинного обучения и искусственного интеллекта. Уметь: Использовать цифровые средства общения для обмена данными и коллаборации на проектах в области машинного обучения и искусственного интеллекта, находить и использовать ресурсы для повышения своих навыков в области машинного обучения и искусственного интеллекта, создавать контент	Задание 1 В Google Colaboratory на выбранном датасете обучить и провести оценку модели машинного обучения. Разместите блокнот и модель в GitHub, добавьте вашу модель в репозиторий и установите ее как зависимость для ваших Jupyter Notebooks Задание 2 Импортировать из GitHub датасет и Jupyter Notebook из GitHub, используйте библиотеки Python для машинного обучения, такие как Scikit-learn или TensorFlow, для создания и обучения модели, разместите блокнот и модель в GitHub

		для обмена знаниями в области машинного обучения и искусственного интеллекта.	
	2. Владеет навыками организации взаимодействия и коммуникации с помощью информационных систем и/или цифровых сервисов и технологий.	Знать: Методы использования цифровых средств общения для обмена данными и коллаборации на проектах в области машинного обучения и искусственного интеллекта, как использовать цифровые средства общения для коммуникации с экспертами в области машинного обучения и искусственного интеллекта для получения консультаций и решений проблем. Уметь: Использовать цифровые средства общения для обмена данными и коллаборации на проектах в области машинного обучения и искусственного интеллекта, использовать цифровые средства общения для коммуникации с экспертами в области машинного обучения и искусственного интеллекта для получения консультаций и решений проблем.	Задание 1 Выполните прогнозирование временных рядов с использованием фреймворка TensorFlow, используя сверточные и рекуррентные нейронные сети (CNN и RNN). Сравните и оцените полученные результаты. Задание 2 В репозитории TensorFlow Hub (содержащий предварительно обученные нейронные сети для решения многих видов задач, в основном в области компьютерного зрения и обработки естественного языка) найдите модули для классификации изображений. Используя технологии <code>tf2_image_retraining</code>, обучите модель в условиях ограниченного набора данных и времени для вычислений.
	3. Осуществляет подбор и применение различных информационно-коммуникационных средств для решения образовательных и профессиональных задач.	Знать: Преимущества и недостатки различных информационно-коммуникационных средств для решения задач в области искусственного интеллекта в зависимости от конкретной прикладной задачи. Уметь: Использовать различные информационно-коммуникационные средства для решения задач в области искусственного интеллекта в зависимости от конкретной прикладной задачи.	Задание 1 Используя библиотеку KerasNLP, решите задачу анализа тональности текста, с предварительно обученным классификатором, выполните тонкую настройку индивидуальной модели. Задание 2 Используя нейросетевую модель LSTM, проведите анализ временных рядов KPI деятельности организации, проведите прогнозирование на несколько периодов.

Примерные вопросы к зачету:

1. Выбрать индивидуальное задание: задачи классификации, регрессии, обработка текста и др. (датасеты с Kaggle <https://www.kaggle.com/datasets?fileType=csv>)
2. В Colab (или используя платформы Data Science and Machine Learning - любую на выбор: Loginom, Knime, RapidMiner) выполнить консолидацию данных кейса, если кейс состоит из нескольких таблиц

3. провести очистку, трансформацию данных (ETL), выполнить preprocessing данных (<https://www.sv-europe.com/crisp-dm-methodology/>), провести EDA с использованием любой библиотеки python D-Tale, pandas-profiling, Sweetviz, AutoViz, Dataprep, dabl, Bamboolib, MitoSheet.
4. обучить и оценить по метрикам качества модели машинного обучения (использовать различные алгоритмы); сравнить полученные результаты.

Примерные вопросы к экзамену:

1. Охарактеризуйте области применения искусственных нейронных сетей.
2. Охарактеризуйте биологический нейрон, структуру и свойства искусственного нейрона.
3. Охарактеризуйте разновидности искусственных нейронов.
4. Охарактеризуйте постановку задачи обучения по прецедентам. Параметры и гиперпараметры. Недообучение и переобучение.
5. Назовите и охарактеризуйте виды активационных функций.
6. Охарактеризуйте классификацию искусственных нейронных сетей и их свойства.
7. Охарактеризуйте принципы работы многослойного персептрона.
8. Охарактеризуйте значение теорема Колмогорова-Арнольда, и следствия из неё.
9. Охарактеризуйте постановку и возможные пути решения задачи обучения искусственных нейронных сетей.
10. Охарактеризуйте обучение с учителем, алгоритм обратного распространения ошибки.
11. Охарактеризуйте обучение без учителя.
12. Охарактеризуйте нейронные сети радиальных базисных функций.
13. Охарактеризуйте вероятностную нейронную сеть.
14. Охарактеризуйте обобщенно-регрессионную нейронную сеть.
15. Охарактеризуйте нейронную сеть Кохонена.
16. Охарактеризуйте нейронные сети встречного распространения.
17. Охарактеризуйте нейронные сети Хопфилда.
18. Охарактеризуйте нейронные сети Хэмминга.
19. Охарактеризуйте двунаправленная ассоциативная память.
20. Охарактеризуйте каскадные искусственные нейронные сети.
21. Охарактеризуйте сети адаптивной резонансной теории.
22. Охарактеризуйте когнитрон и неоккогнитрон.
23. Рекуррентные нейронные сети.
24. Охарактеризуйте нейронные сети GRU, LSTM.
25. Охарактеризуйте двусторонние рекуррентные нейронные сети.
26. Охарактеризуйте автоэнкодеры, вариационные автоэнкодеры.
27. Охарактеризуйте генеративно-сопоставительные модели.
28. Охарактеризуйте архитектуру seq2seq.

29. Охарактеризуйте обучение с подкреплением.
30. Охарактеризуйте архитектуру w2v.
31. Охарактеризуйте механизмы внимания.
32. Охарактеризуйте архитектуру Transformer.

Примерные практические задания к экзамену:

1. В платформе Colab Laboratory выполнить классификацию рукописных цифр с помощью нейронных сетей:
 - Обучите линейную модель и нейронную сеть для классификации рукописных цифр из классического набора данных [MNIST](https://www.yale.edu/mpeg/ISIPV/mnist/). Повысьте точность линейной модели путем экспериментирования с гиперпараметрами размера пакета, скорости обучения и числа шагов (batch size, learning rate, steps).
 - Сравните производительность моделей классификации линейных и нейронных сетей. Найдите комбинацию параметров нейронной сети, которая дает точность распознавания 0,95.
 - Визуализировать вес скрытого слоя нейронной сети
2. Используйте LinearRegressor во фреймворке TensorFlow, предскажите среднюю цену на жилье в Калифорнии.
 - Оцените точность прогнозов модели, используя среднеквадратическую ошибку (RMSE)
 - Повысьте точность модели, настроив ее гиперпараметры.
3. В среде Colab Laboratory решить задачу двоичной классификации с выбранным на <https://www.kaggle.com/datasets> кейсом (например, [набором данных по жилью в Калифорнии](https://www.kaggle.com/datasets)), применяя глубокое и классическое машинное обучение. Учитывая признаки (features), описывающие городской квартал, постройте прогноз (например, является ли городской квартал дорогим), сравните полученные результаты, оцените важность признаков.
4. В среде Colab Laboratory повысить производительность нейронной сети путем нормализации функций и применения различных алгоритмов оптимизации:
 - нормализуйте входные данные в масштабе -1, 1. Проведите обучение и оценку нейронной сети на новых нормализованных данных.
 - используйте оптимизаторы Adagrad и Adam и сравните производительность.
5. В среде Colab Laboratory решить задачу классификации текста с выбранным на <https://www.kaggle.com/datasets?tags=13204-NLP> кейсом, оценить

Пример экзаменационного билета:

1. Охарактеризуйте принципы работы многослойного персептрона (15 баллов).
2. В среде Colab Laboratory повысить производительность нейронной сети путем нормализации функций и применения различных алгоритмов оптимизации:
 - нормализуйте входные данные в масштабе -1, 1. Проведите обучение и оценку нейронной сети на новых нормализованных данных.
 - используйте оптимизаторы Adagrad и Adam и сравните производительность (45 баллов).

Методические материалы, определяющие процедуры оценивания знаний, умений

Приказ от 23.03.2017 №0557/о «Об утверждении Положения о проведении текущего контроля успеваемости и промежуточной аттестации обучающихся по программам бакалавриата и магистратуры в Финансовом университете».

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

Нормативно-правовые акты:

1. Федеральный Закон Российской Федерации «Об информации, информационных технологиях и о защите информации» № 149-ФЗ от 13.07.2015 г.
2. Паспорт национальной программы «Цифровая экономика Российской Федерации» (утв. президиумом Совета при Президенте РФ по стратегическому развитию и национальным проектам, протокол от 24.12.2018 N 16).
3. Государственная программа Российской Федерации "Информационное общество (2011-2020 годы)" (в ред. Постановления Правительства РФ от 18.05.2011 N 399).

а) основная:

1. Плас, Дж. В. Python для сложных задач: наука о данных и машинное обучение : практическое руководство : пер. с англ. / Дж. В. Плас. – Санкт-Петербург : Питер, 2021. – 576 с. – (Бестселлеры O'Reilly). – ISBN 978-5-4461-0914-2. – ЭБС Znanium.com. – URL: <https://znanium.com/catalog/product/1739601> (дата обращения: 13.12.2023). – Текст : электронный.
2. Силен, Д. Основы Data Science и Big Data. Python и наука о данных : практическое руководство / Д. Силен, А. Мейсман, М. Али. – Санкт-Петербург : Питер, 2018. – 336 с. – (Библиотека программиста). – ISBN 978-5-496-02517-1. – ЭБС Znanium.com. – URL: <https://znanium.com/catalog/product/1766398> (дата обращения: 13.12.2023). – Текст : электронный.

б) дополнительная:

3. Лесковец, Ю. Анализ больших наборов данных : практическое руководство / Д. Дж. Ульман, Ю. Лесковец, А. Раджараман ; пер. с англ. А. А. Слинкина. – 2-е изд. – Москва : ДМК Пресс, 2023. – 500 с. – ISBN 978-5-89818-304-2. – ЭБС Znanium.com. – URL: <https://znanium.com/catalog/product/2102592> (дата обращения: 13.12.2023). – Текст : электронный.
4. Дадян, Э. Г. Данные: хранение и обработка : учебник / Э. Г. Дадян ; Финуниверситет. — Москва : ИНФРА-М, 2021. – 205 с. – (Высшее образование: Бакалавриат). – ISBN 978-5-16-016447-2. – ЭБС Znanium.com. – URL: <https://znanium.com/catalog/product/1149101> (дата обращения: 13.12.2023).

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

1. <https://cloud.google.com/bigquery/docs/sandbox> – страница BigQuery sandbox.
2. <https://www.anaconda.com/> – страница загрузки Anaconda.
3. <https://cloud.yandex.ru/services/datalens> – сервис визуализации и анализа данных Яндекс.
4. Портал открытых данных Российской Федерации <https://data.gov.ru>
5. <https://cloud.yandex.ru/training/corpplatform> - практический курс «Построение корпоративной аналитической платформы»
6. <https://practicum.yandex.ru/ycloud/> - бесплатный курс «Инженер облачных сервисов»
7. [https://rise.articulate.com/share/BtQjK0gEy1lktRK6q2hPZ5KnRDJhB8k#/?](https://rise.articulate.com/share/BtQjK0gEy1lktRK6q2hPZ5KnRDJhB8k#/) - Ростелеком бизнес. Платформа управления данными
8. Электронная библиотека Финансового университета (ЭБ) <http://elib.fa.ru/>
9. Электронно-библиотечная система BOOK.RU <http://www.book.ru>
10. Электронно-библиотечная система Znanium <http://www.znanium.com>

10. Методические указания для обучающихся по освоению дисциплины

Студентам необходимо руководствоваться «Методическими рекомендациями по планированию и организации внеаудиторной самостоятельной работы по образовательным программам бакалавриата и магистратуры в Финансовом университете» (Приказ ректора № 1040_о от 11.05.2021) и данной рабочей программой дисциплины.

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. ОС Astr Linux,
2. LibreOffice

3. Антивирус Kaspersky

11.2 Современные профессиональные демонстрационные и информационные справочные системы:

1. Консультант Плюс.

11.3 Сертифицированные программные и аппаратные средства защиты информации:

Не предусмотрены.

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

Помещения для проведения лекций, семинарских занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы.