

**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

21.04. 2023 г. Е.А. Каменева

П.А. Сахнюк, Х.Х. Кучмезов

Технологии продвинутой аналитики

Рабочая программа дисциплины
для студентов, обучающихся по направлению подготовки
38.03.05 - Бизнес-информатика
ОП «Цифровая трансформация управления бизнесом»
профили: «Технологии цифровых бизнес-моделей»,
«ИТ-менеджмент в бизнесе»

*Рекомендовано Ученым советом Факультета информационных
технологий и анализа больших данных
(протокол № 31 от 18.04.2023 г.)*

*Одобрено Советом учебно-научного Департамента бизнес-
информатики
(протокол № 7 от 28.03.2023 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.....	4
4. Объем дисциплины в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся.....	5
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.....	7
5.3. Содержание семинаров, практических занятий.....	8
6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине.....	9
6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы.....	9
6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю.....	10
7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине.....	11
8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины.....	15
9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины.....	16
10. Методические указания для обучающихся по освоению дисциплины.....	17
11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем.....	17
12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине.....	17

1. Наименование дисциплины

«Технологии продвинутой аналитики».

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
ПКН-3	Способность применять аналитические системы и консультировать по вопросам разработки и развития аналитических систем работы с данными	1. Применять аналитические системы работы с данными	Знать: современные приемы использования аналитических систем для работы с данными Уметь: использования современные аналитические системы для работы с данными
		2. Проводит анализ рынка аналитических систем работы с данными	Знать: теоретические основы анализа рынка аналитических систем работы с данными Уметь: анализировать рынок аналитических систем для работы с данными
		3. Консультирует по вопросам применения аналитических систем работы с данными	Знать: специфические особенности функционала аналитических систем работы с данными Уметь: настраивает аналитические системы в соответствии с требованиями бизнес-заказчика
ПКП-3 <i>Профиль «ИТ-менеджмент в бизнесе»</i>	Способность предлагать различные варианты инфраструктурных решений для поддержки ИТ/ИС	1. Анализирует текущий уровень инфраструктурных решений предприятия/организации	Знать: Современное состояние ИТ инфраструктурные организации. Уметь: анализировать текущий уровень инфраструктурных решений организации.
		2. Формирует и обосновывает варианты технологического	Знать: эффективные методы создания и анализа технологического слоя

		слоя архитектуры предприятия/организации	архитектуры предприятия Уметь: создавать, формировать и модернизировать, обосновывать варианты технологического слоя архитектуры предприятия
ПКП-3 <i>Профиль «Технологии цифровых бизнес-моделей»</i>	Способность управлять процессами цифровой трансформации бизнеса	1.Проектирует оптимальную модель управления процессом цифровой трансформации, опираясь на результаты анализа готовности компании к цифровым преобразованиям.	Знать: эффективный набор управленческих инструментов и подходов на базе лучших мировых практик цифровой трансформации Уметь: применять технологический подход к цифровой трансформации, как совокупность технологий, внедрение и реализация которых обеспечивают компании переход в новое цифровое пространство
		2.Определяет цели и ожидаемые результаты трансформации бизнеса, необходимые ресурсы для ее реализации и методы управления ключевыми рисками.	Знать: цели и ожидаемые результаты трансформации бизнеса, основные методы управления рисками, методы управления ключевыми рисками Уметь: определять необходимые ресурсы для реализации цифровой трансформации бизнеса, использовать возможности для совершенствования и снижения рисков

3. Место дисциплины в структуре образовательной программы

Дисциплина «Технологии продвинутой аналитики» относится к циклу профиля (элективный) части, формируемой участниками образовательных отношений по направлению подготовки 38.03.05 - Бизнес-информатика ОП «Цифровая трансформация управления бизнесом».

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 7 (в часах)
Общая трудоемкость дисциплины	3 ЗЕ./108 ч.	108
<i>Контактная работа - Аудиторные занятия</i>	30	30
<i>Лекции</i>	14	14
<i>Семинары, практические занятия</i>	16	16
Самостоятельная работа	78	78
Вид текущего контроля	Контрольная работа	Контрольная работа
Вид промежуточной аттестации	Зачет	Зачет

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. Deepnote

Бесплатный онлайн-блокнот Deepnote для специалистов по данным, фокусирующийся в основном на совместном использовании в реальном времени и абстрагировании всех понятий, усложняющих работу аналитика, таких как настройка окружения и инфраструктуры. Используемые языки: Jupyter-совместимый, языки программирования Jupyter (например, Python, R), SQL. Коннекторы: подключения к библиотекам Jupyter (например, SQLAlchemy, psycopg2), к хранилищам данных и озерам данных (AWS, GCP и т.д.), базам данных (PostgreSQL, MongoDB и т.д.), предоставляемое файловое хранилище. Визуализация данных Deepnote: Jupyter (например, библиотеки Matplotlib, Altair, Plotly), пользовательский интерфейс для построения диаграмм.

Машинное обучение – поддерживает интеграции: встроенная интеграции с MindsDB, интеграции с Weights & Biases, Comet.ml, Neptune.ai.

MindsDB привносит машинное обучение в базы данных, используя концепцию AI Tables — это модели машинного обучения, хранящиеся в виде виртуальных таблиц внутри базы данных. Позволяют выполнять прогнозы временных рядов, регрессии и классификации в своей базе данных и получать результаты, запрашивая таблицу AI с помощью простых операторов SQL. Weights & Biases помогает отслеживать проекты машинного обучения. Используйте его для регистрации гиперпараметров и вывода метрик из ваших прогонов или для визуализации и сравнения результатов. Comet позволяет специалистам по данным и командам отслеживать, сравнивать, объяснять и оптимизировать эксперименты и модели на протяжении всего жизненного цикла модели. Neptune – это легкий

инструмент управления экспериментами, который помогает отслеживать эксперименты с машинным обучением.

Тема 2. Jetbrains Datalore

Jetbrains Datalore – это мощная онлайн-среда для Jupyter-ноутбуков, разработанная в JetBrains: использование интеллектуальной помощи по написанию кода для Python в онлайн-блокнотах Jupyter, запуск кода на мощных процессорах и графических процессорах, сотрудничество в режиме реального времени и легкость в возможностях делиться результатами. Анализ данных с помощью умного кодирования и автоматизации без кода. Создание интерактивных приложений для работы с данными, созданными из записных книжек всего за несколько кликов.

Функции: Jupyter-совместимый; языки программирования Jupyter (например, Python, R); подключение к библиотекам Jupyter (например, SQLAlchemy, psycopg2), базам данных (MySQL, PostgreSQL), хранилищам данных и озерам данных (AWS, GCP и т.д.), предоставляемое файловое хранилище; визуализация данных Jupyter (например, Matplotlib, Altair, Plotly) и пользовательский интерфейс для построения диаграмм; полная реактивность в реальном времени; встроен планировщик записной книжки.

Тема 3. Платформа Mode

Mode – платформа, объединяющая команды по обработке данных и бизнес-группы вокруг данных для достижения бизнес-результатов. В Mode SQL, R, Python и визуализация данных связаны друг с другом, чтобы помочь вам получать ценную информацию: анализ данных с помощью SQL, используя блокноты R или Python (нативные блокноты R и Python). В отличие от многих других инструментов бизнес-аналитики, в Mode есть все функции и возможности, необходимые для решения вопросов расширенной аналитики, – и все это на одной платформе. Переключение с SQL на полностью оборудованную среду для ноутбуков Python и R одним кликом мыши.

Mode не совместима с Jupyter, языки программирования: SQL, используя блокноты R или Python. Подключение к базам данных (MariaDB, PostgreSQL и др.), хранилищам данных и озерам данных (Amazon Redshift, Google BigQuery и др.). Помимо библиотек Python использует пользовательский интерфейс для построения диаграмм. Интерактивные информационные панели без кода, которые может создать каждый: создание группой аналитиков данных дашбордов в рекордно короткие сроки, а бизнес-группы могут использовать их в качестве отправной точки для самообслуживания.

Тема 4. Библиотеки python для анализа больших объемов структурированных данных: Polar, Vaex, Dask, Modin, Ray.

Библиотека Polars использует эффективные алгоритмы распараллеливания и кэширования для ускорения выполнения аналитических

задач: уменьшение количества избыточных копий; эффективное кэширование памяти; сведение конфликтов при параллелизме к минимуму; реализована на Rust, а не на Python (Rust намного лучше реализует параллелизм, нежели Python).

Dask: планировщик низкого уровня и частичная замена Pandas высокого уровня, ориентированный на выполнение кода на вычислительных кластерах. Ray: низкоуровневый фреймворк для распараллеливания кода Python между процессорами или кластерами. Modin: замена Pandas на базе Dask или Ray. Vaex: частичная замена Pandas, использующая ленивую оценку и сопоставление памяти, чтобы позволить разработчикам работать с большими наборами данных на стандартных машинах.

Modin с Ray в качестве бэкенда. Установив их, можно добиться значительное преимущество, изменив всего одну строку («import pandas as pd» на «import modin.pandas as pd»). В отличие от других инструментов, Modin стремится достичь полной совместимости с Pandas. Dask, более крупный и, следовательно, более сложный проект. Dask также предоставляет Dask.dataframe, библиотеку более высокого уровня, подобную Pandas, которая помогает справиться с наборами данных вне ядра. Vaex, предназначенный для работы с большими объемами данных на стандартном ноутбуке. Его замена Pandas охватывает некоторые API Pandas, но больше ориентирована на исследование и визуализацию.

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		
1	Тема 1. Deepnote	28	8	4	4	20	Выполнение индивидуальных заданий
2	Тема 2. JetBrains Datalore	28	8	4	4	20	Выполнение индивидуальных заданий
3	Тема 3. Платформа Mode	24	6	2	4	18	Выполнение индивидуальных заданий
4	Тема 4. Библиотеки python для анализа больших объемов структурированных данных: Polar, Vaex, Dask, Modin, Ray.	28	8	4	4	20	Выполнение индивидуальных заданий
	В целом по дисциплине	108	30	14	16	78	Контрольная работа

Итого в%:		28	47	53	72	
-----------	--	----	----	----	----	--

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. Deepnote	1. Разработка сценариев ABC, XYZ, ABC-XYZ анализов в Deepnote 2. Визуализация результатов аналитики: в библиотеки Altair, Plotly. Основная литература: 1,2 Дополнительная литература: 1,2,3	Выполнение и защита практических заданий
Тема 2. JetBrains Datalore	1. Разработка сценариев RFM-анализа в Datalore 2. Создание интерактивных приложений на основе анализа данных. Основная литература: 1,2 Дополнительная литература: 1,2,3	Выполнение и защита практических заданий
Тема 3. Платформа Mode	1. Разработка сценариев когортного анализа в Datalore 2. Создание интерактивных информационных панелей на основе анализа данных. Основная литература: 1,2 Дополнительная литература: 1,2,3	Выполнение и защита практических заданий
Тема 4. Библиотеки python для анализа больших объемов структурированных данных: Polar, Vaex, Dask, Modin, Ray.	1. Базовая аналитика и описательные статистики на больших объемах структурированных данных с использованием библиотек Polar, Vaex, Dask, Modin, Ray. 2. Визуальная аналитика больших объемах структурированных данных 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. Deepnote	Машинное обучение в Deepnote – интеграции: интеграции с Weights & Biases, Comet.ml, Neptune.ai.	Изучение методических материалов по теме в электронном виде и рекомендуемых разделов основной и дополнительной литературы, интернет – источников. Подготовка к семинарам
Тема 2. JetBrains Datalore	Создание интерактивных приложений на основе ABC-XYZ анализов в JetBrains Datalore	Изучение методических материалов по теме в электронном виде и рекомендуемых разделов основной и дополнительной литературы, интернет – источников. Подготовка к семинарам
Тема 3. Платформа Mode	Создание интерактивных информационных панелей на основе совместного ABC-XYZ, RFF и когортного анализа	Изучение методических материалов по теме в электронном виде и рекомендуемых разделов основной и дополнительной литературы, интернет – источников. Подготовка к семинарам
Тема 4. Библиотеки python для анализа больших объемов структурированных данных: Polar, Vaex, Dask, Modin, Ray.	Машинное обучение в Dask	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:

Студенту в начале семестра предлагается выбрать кейс, на основе которого разрабатываться аналитический проект. Для этого можно использовать датасеты с Kaggle <https://www.kaggle.com/datasets?search=customer>, содержащие клиентские данные (всего 2466 различных Datasets).

Например:

1. Customer Personality Analysis.
2. Telco Customer Churn.
3. Shop Customer Data.
4. Customer Segmentation.
5. Mall Customer Segmentation Data.
6. Credit Card customers.
7. Starbucks Customer Survey.
8. Airlines Customer satisfaction.
9. Loan Prediction Based on Customer Behavior.
10. Brazilian E-Commerce Public Dataset by Olist.

Задание:

1. Импортировать данные в BigQuery. Обогащать датасет, проводя клиентскую аналитику: выполнить ABC-XYZ, RFM, когортный -анализ, используя библиотеки Python.
2. Создать, обучить, оценить по метрикам качества, объяснить модели машинного обучения библиотеки Python; создать интерактивный отчет (дашборд) в Datalore (или Mode, Deepnote).

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
ПКН-3 Способность применять аналитические системы и консультировать по вопросам разработки и развития аналитических систем работы с данными	1. Применять аналитические системы работы с данными	Знать: современные приемы использования аналитических системы для работы с данными Уметь: использования современные аналитические системы для работы с данными	Задание 1 Осуществите проектирование дашбордов, отражающих основные перспективы системы сбалансированных показателей эффективности и мониторинг KPI организации, используя Datalore (Mode, Deepnote). Задание 2 Используя библиотеку Dask проведите прогнозирование временных рядов финансовых показателей, проведите интерпретацию полученных результатов.
	2. проводить анализ рынка аналитических систем работы с данными	Знать: теоретические основы анализа рынка аналитических систем работы с данными Уметь: анализировать рынок аналитических систем для работы с данными	Задание 1 Проведите анализ текущего состояния рынка аналитических систем работы с данными, рассматривая основные компании на российском рынке. Охарактеризуйте тенденции и направления развития рынка: интеграция BI-систем с инструментами машинного обучения, разработка инструментов предиктивной аналитики, внедрение Self-Service BI. Задание 2 Проведите анализ текущего состояния мирового рынка BI-систем (Business Intelligence).

			Охарактеризуйте тенденции и направления развития рынка: интеграция BI-систем с инструментами машинного обучения, разработка инструментов предиктивной аналитики, внедрение Self-Service BI.
	3. Консультирует по вопросам применения аналитических систем работы с данными	Знать: специфические особенности функционала аналитических систем работы с данными Уметь: настраивает аналитические системы в соответствии с требованиями бизнес-заказчика	Задание 1 Для выбранного предприятия определите области, в которых могут быть использованы аналитические системы. Укажите возможные направления их внедрения и эффект от их использования. Задание 2 Для выбранного предприятия проведите консультацию заказчика по вопросам разработки ETL-процедур (извлечение, трансформация и загрузка данных), проектирования дизайна интерактивных панелей в соответствии со сценариями анализа, специальными требованиями заказчика (корпоративный дизайн, brandbook), визуальными возможностями выбранной аналитической системы.
ПКП-3 Способность предлагать различные варианты инфраструктурных решений для поддержки ИТ/ИС <i>Профиль «ИТ-менеджмент в бизнесе»</i>	1. Анализирует текущий уровень инфраструктурных решений	Знать: современное состояние ИТ инфраструктурные организации. Уметь: анализировать текущий уровень инфраструктурных решений организации.	Задание 1 Для выбранного предприятия выявите необходимых изменений в бизнес-процессах, определите мероприятия адаптации учетных систем и изменение форм отчетности для внедрения систем бизнес-аналитики, разработайте системы ключевых показателей эффективности (KPI). Разработайте предложения для заказчика по выбору направлений изменений ИТ инфраструктуры организации. Задание 2 Для выбранного предприятия определите состав информации, который следует анализировать для повышения эффективности

			бизнеса, выявите возможности по получению требуемых данных, проведите проектирование целевой архитектуры ИТ инфраструктуры организации. Разработайте предложения для заказчика по выбору направлений изменений ИТ инфраструктуры организации.
	2. Формирует и обосновывает варианты технологического слоя архитектуры предприятия/организации	Знать: эффективные методы создания и анализа технологического слоя архитектуры предприятия Уметь: создавать, формировать и модернизировать, обосновывать варианты технологического слоя архитектуры предприятия	Задание 1 Для выбранного предприятия определите мероприятия адаптации технологического слоя архитектуры предприятия для внедрения аналитической системы. Сформируйте и обоснуйте предложения для заказчика по выбору варианта технологического слоя архитектуры предприятия. Задание 2 Для выбранного предприятия определите необходимый технологический состав архитектуры предприятия. Разработайте предложения для заказчика по выбору направлений модернизации технологического слоя, для обеспечения информационной поддержки деятельности предприятия.
ПКП-3 Способность управлять процессами цифровой трансформации бизнеса <i>Профиль «Технологии цифровых бизнес-моделей»</i>	1.Проектирует оптимальную модель управления процессом цифровой трансформации, опираясь на результаты анализа готовности компании к цифровым преобразованиям.	Знать: эффективный набор управленческих инструментов и подходов на базе лучших мировых практик цифровой трансформации Уметь: применять технологический подход к цифровой трансформации, как совокупность технологий, внедрение и реализация которых	Задание 1 Для выбранного предприятия определите мероприятия по облачному развертыванию аналитической системы, используя современные технические средства стека технологий платформ Deepnote (Datalore, Mode). Задание 2 Используя инструменты и технологии платформ Deepnote (Datalore, Mode) осуществите консолидацию данных из разных источников, очистку и предобработку данных;

		обеспечивают компании переход в новое цифровое пространство	используя стандартные визуализации создайте аналитический интерактивный отчет, выявите инсайты. Разместите интерактивное приложение в облаке, организуйте совместную работу внутри команды. Предложите вариант изменения бизнес-модели предприятия.
	2. Определяет цели и ожидаемые результаты трансформации и бизнеса, необходимые ресурсы для ее реализации и методы управления ключевыми рисками.	Знать: цели и ожидаемые результаты трансформации бизнеса, основные методы управления рисками, методы управления ключевыми рисками Уметь: определять необходимые ресурсы для реализации цифровой трансформации бизнеса, использовать возможности для совершенствования и снижения рисков	Задание 1 Для выбранного предприятия определите цели и задачи бизнеса, которые можно решить цифровой трансформацией. Разработайте предложения для заказчика по выбору направлений снижения рисков и негативных последствий, а также управляющие воздействия по преодолению их отрицательного влияния. Задание 2 Создайте панели мониторинга KPI финансовых показателей, используя технологии платформ Deepnote (Datalore, Mode). Используя встроенные инструменты расширенной аналитики проведите прогнозирование временных рядов финансовых показателей, проведите интерпретацию полученных результатов.

Примерные вопросы к зачету:

1. Охарактеризуйте основные возможности онлайн-блокнота Deepnote.
2. Охарактеризуйте ограничения онлайн-блокнота Deepnote.
3. Охарактеризуйте основные возможности онлайн-среды JetBrains Datalore.
4. Охарактеризуйте ограничения онлайн-среды JetBrains Datalore.
5. Охарактеризуйте основные возможности платформы Mode.
6. Охарактеризуйте ограничения платформы Mode.
7. Охарактеризуйте основные возможности библиотеки Polar.
8. Охарактеризуйте ограничения библиотеки Polar.
9. Охарактеризуйте основные возможности библиотеки Vaex.
10. Охарактеризуйте ограничения библиотеки Vaex.
11. Охарактеризуйте основные возможности библиотеки Dask.

12. Охарактеризуйте ограничения библиотеки Dask.
13. Охарактеризуйте основные возможности библиотеки Modin.
14. Охарактеризуйте ограничения библиотеки Modin.
15. Охарактеризуйте основные возможности библиотеки Ray.
16. Охарактеризуйте ограничения библиотеки Ray.

Примерные практические задания к зачету:

1. Разработка сценариев ABC-XYZ, RFM, когортного - анализ в выбранной среде Deenote (или Datalore, Mode).
2. Визуализация результатов аналитики: в библиотеки Altair, Plotly, а также встроенными средствами платформ Deenote (или Datalore, Mode).
3. Базовая аналитика и описательные статистики на больших объемах структурированных данных с использованием библиотек Polar, Vaex, Dask, Modin, Ray.
4. Визуальная аналитика больших объемах структурированных данных с использованием библиотек Polar, Vaex, Dask, Modin, Ray.

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 учебной 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. Бринк, Х. Машинное обучение / Х. Бринк, Ричарде Дж., М. Феверолф. – СПб.: Питер, 2017. – 336 с.
2. де Прадо М. «Машинное обучение: алгоритмы для бизнеса», Санкт-Петербург: Издательский дом «Питер». – 2019. – 432 с.
3. Воронина, В. В. Теория и практика машинного обучения : учеб. пособие / В. В. Воронина. — Ульяновск : УлГТУ, 2017. — 290 с. — ISBN 978-5-

- 9795-1712-4. — ЭБС Лань. — URL: <https://e.lanbook.com/book/165053> (дата обращения: 15.05.2023). — Текст : электронный.
4. Борзунов, С. В. Языки программирования. Python: решение сложных задач / С. В. Борзунов, С. Д. Кургалин. — Санкт-Петербург : Лань, 2023. — 192 с. — ISBN 978-5-507-45923-0. — ЭБС Лань. — URL: <https://e.lanbook.com/book/319394> (дата обращения: 15.05.2023). — Текст : электронный.
 5. Протоdjяконов, А. В. Алгоритмы Data Science и их практическая реализация на Python : учеб. пособие / А. В. Протоdjяконов, П. А. Пылов, В. Е. Садовников. — Москва : Вологда : Инфра-Инженерия, 2022. — 392 с. — ISBN 978-5-9729-1006-9. — ЭБС Znanium.com. — URL: <https://znanium.com/catalog/product/1902689> (дата обращения: 15.05.2023). — Текст : электронный.

б) дополнительная:

6. Марц, Н. Большие данные = Big Data. Принципы и практика построения масштабируемых систем обработки данных в реальном времени : пер. с англ. / Н. Марц, Д. Уоррен. — Москва : Вильямс, 2016. — 368 с. — ISBN 978-5-8459-2075-1. — Текст : непосредственный.
7. Гуриков, С. Р. Основы алгоритмизации и программирования на Python : учеб. пособие / С. Р. Гуриков. — Москва : ИНФРА-М, 2023. — 343 с. — (Высшее образование: Бакалавриат). — ISBN 978-5-16-017142-5. — ЭБС Znanium.com. — URL: <https://znanium.com/catalog/product/1913856> (дата обращения: 15.05.2023). — Текст : электронный.
8. Келлехер, Д. Наука о данных : базовый курс : учеб. пособие : пер. с англ. / Д. Келлехер, Б. Тирни ; науч. ред. З. Мамедьяров. — Москва : Альпина Паблишер, 2020. — 224 с. : схем., табл. — ISBN 978-5-9614-3170-4. — ЭБС Университетская библиотека онлайн. — URL: <https://biblioclub.ru/index.php?page=book&id=598235> (дата обращения: 15.05.2023). — Текст : электронный.
9. Дадян, Э. Г. Данные: хранение и обработка : учебник / Э. Г. Дадян. — 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: 15.05.2023). — Текст : электронный.

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/BtQjK0gEy1lktRKR6q2hPZ5KnRDJhB8k#/#/](https://rise.articulate.com/share/BtQjK0gEy1lktRKR6q2hPZ5KnRDJhB8k#/) - Ростелеком бизнес. Платформа управления данными

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.