

**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
25.01. 2023 г.

П.А. Сахнюк

Технологии продвинутой аналитики

Рабочая программа дисциплины
для студентов, обучающихся по направлению подготовки
38.04.01 – Экономика,
направленность программы:
«Финансовый анализ и оценка инвестиционных решений»
«Анализ и стратегический менеджмент в бизнесе»

*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.....	5
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.....	6
5.1. Содержание дисциплины.....	6
5.2. Учебно-тематический план.....	7
5.3. Содержание семинаров, практических занятий.....	7
6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине.....	8
6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы.....	8
6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю.....	9
7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине.....	9
8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины.....	14
9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины.....	15
10. Методические указания для обучающихся по освоению дисциплины.....	15
11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем.....	15
12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине.....	15

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-6 <i>«Финансовый анализ и оценка инвестиционных решений»</i>	Способность анализировать и прогнозировать основные социально-экономические показатели, предлагать стратегические направления экономического развития на микро-, мезо- и макроуровнях	1. Применяет методический инструментарий системного анализа и моделирования экономических процессов для обоснования внедрения инновационных разработок с целью получения конкурентных преимуществ и обеспечения опережающего роста на новых и развивающихся рынках	Знать: методы решения задач Data Mining, применимых для анализа и прогноза основных социально-экономических показателей Уметь: применять методы машинного обучения для анализа и моделирования экономических процессов для обоснования внедрения инновационных разработок
		2. Обосновывает перспективы изменений основных социально-экономических показателей и стратегические направления экономического развития на микро-, мезо- и макроуровнях	Знать: метрики качества и способы объяснения моделей машинного обучения, применяемых для анализа и прогнозирования основных социально-экономических показателей Уметь: обосновывать принимаемые решения, полученные на основе применения методов машинного обучения на различных уровнях

ПК-6 <i>«Финансовый анализ и оценка инвестиционных решений»</i>	Способность применять методы интеллектуального анализа данных к решению прикладных задач, связанных с оказанием финансовых услуг	1. Владеет основными методами интеллектуального анализа данных и машинного обучения, используемые в экономике и финансах.	Знать: основные методы интеллектуального анализа данных и машинного обучения, используемые в экономике и финансах Уметь: применять основные методы интеллектуального анализа данных и машинного обучения в экономике и финансах
		2. Владеет профессиональной терминологией в области больших данных и машинного обучения	Знать: основные теоретические положения в области машинного обучения и анализа данных Уметь: использовать открытую документацию по платформам, фреймворкам и библиотекам машинного обучения и анализа данных
		3. Владеет современными пакетами прикладных программ, реализующими методы обработки больших данных и машинного обучения при оказании финансовых услуг	Знать: спектр доступных библиотек, фреймворков и платформ машинного обучения и аналитики больших данных, применяемых в экономике и финансах Уметь: применять пакеты прикладных программ, библиотеки, фреймворки и платформы машинного обучения и аналитики больших данных при оказании финансовых услуг
		4. Применяет модели машинного обучения для решения практических задач в области предоставления финансовых услуг	Знать: задачи и методы Data Mining, методiku проведения кросс-индустриального стандарта Data Mining (CRISP DM) Уметь: использовать библиотеки python для решения задач машинного обучения, включая получение данных, их предобработку, создание и обучение модели, тюнинг модели, оценку и объяснение модели, деплой модели с использованием NumPy, Pandas, Scikit-learn, применять фреймворки автоматизированного машинного обучения.

		5. Обладает навыками анализа и обработки финансовой информации, необходимой для решения задач в области финансового анализа, кредитного скоринга экономического прогнозирования	Знать: особенности различных доступных библиотек, фреймворков и платформ машинного обучения для решения задач в области финансового анализа, кредитного скоринга экономического прогнозирования Уметь: применять различные библиотеки, фреймворки и платформы машинного обучения для решения задач в области финансового анализа, кредитного скоринга экономического прогнозирования
ПК-2 <i>«Анализ и стратегический менеджмент в бизнесе»</i>	Способность применять современные инструменты анализа, в том числе, с использованием интеллектуальных информационно-аналитических систем	1. Проводит самостоятельные исследования бизнес-процессов в соответствии с разработанной производственной программой с использованием современных информационно-аналитических систем.	Знать: основные методы исследования бизнес-процессов с использованием современных информационно-аналитических систем. Уметь: применять основные методы исследования бизнес-процессов в экономике и финансах
		Демонстрирует владение методами сбора, анализа и обработки данных для принятия управленческих решений при моделировании бизнес-процессов	Знать: методологические подходы и методы сбора, анализа и обработки данных для принятия Уметь: моделировать бизнес-процессы на основе методов сбора, анализа и обработки данных для принятия управленческих решений

3. Место дисциплины в структуре образовательной программы

Дисциплина «Технологии продвинутой аналитики» относится модулю дисциплин по выбору, углубляющих освоение программы магистратуры, части, формируемой участниками образовательных отношений направленности программы: «Финансовый анализ и оценка инвестиционных решений» и дисциплиной модуля направленности программы магистратуры «Анализ и стратегический менеджмент в бизнесе» по направлению подготовки 38.04.01 – Экономика.

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

«Финансовый анализ и оценка инвестиционных решений»//«Анализ и стратегический менеджмент в бизнесе»

таблица 2

Вид учебной работы по дисциплине	Всего (в з/ед. и часах)	Модуль 5/6 (в часах)
Общая трудоемкость дисциплины	3 ЗЕ./108 ч.	108
<i>Контактная работа - Аудиторные занятия</i>	32/40	32/40
<i>Лекции</i>	8/8	8/8
<i>Семинары, практические занятия</i>	24/32	24/32
<i>Самостоятельная работа</i>	76/68	76/68
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. 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): PyCaret, AutoGluon, LAMA.

Машинное обучение и технологии на базе искусственного интеллекта (ИИ) в платформах бизнес-аналитики и науки о данных.

Тема 2. Библиотеки python для анализа больших объемов структурированных данных: Polar, Vaex, Dask, Modin, Ray.

Библиотека Polars - эффективные алгоритмы распараллеливания и кэширования для ускорения выполнения аналитических задач: уменьшение количества избыточных копий; эффективное кэширование памяти; сведение конфликтов при параллелизме к минимуму; реализация 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 ленивые вычисления, чтобы позволить работать с большими наборами данных на стандартных компьютерах. Его замена Pandas охватывает некоторые API Pandas, но больше ориентирована на исследование данных и визуализацию.

5.2. Учебно-тематический план

«Финансовый анализ и оценка инвестиционных решений»// «Анализ и стратегический менеджмент в бизнесе»

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. Машинное обучение на python	54/54	16/20	4/4	12/16	38/34	Выполнение индивиду- альных заданий
2	Тема 2. Библиотеки python для анализа больших объемов структурированных данных: Polar, Vaex, Dask, Modin, Ray	54/54	16/20	4/4	12/16	38/34	Выполнение индивиду- альных заданий
	В целом по дисциплине	108	32/40	8/8	24/32	76/68	Контрольная работа
Итого в%:			30/37	25/20	75/80	70/63	

*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.	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	Выполнение и защита практических заданий

	временных рядов. 6. Основные фреймворки машинного обучения с функцией AutoML. 7. Разработка рекомендательной системы. 8. Применение технологий машинного обучения и искусственного интеллекта в платформах BI и Data Science. Основная литература: 1-5 Дополнительная литература: 6-11	
Тема 2. 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 3. Машинное обучение на больших объемах данных. Основная литература: 1-5 Дополнительная литература: 6-11	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	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. Библиотеки python для анализа больших объемов структурированных данных: Polar, Vaex, Dask, Modin, Ray	Машинное обучение с Dask	Изучение методических материалов по теме в электронном виде и рекомендуемых разделов основной и дополнительной

Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
		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. Провести исследовательский анализ данных: правильно прочитать данные, выполнить предварительную обработку данных (preprocessing), рассмотреть описательную статистику, визуализировать данные, выдвинуть предварительные гипотезы.
2. Обогащать датасет, проводя клиентскую аналитику: выполнить ABC-XYZ (RFM, когортный) -анализ, используя библиотеки Python.
3. Создать, обучить, оценить по метрикам качества, объяснить модели машинного обучения библиотеки Python, создать прогноз на новых данных.

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
«Финансовый анализ и оценка инвестиционных решений»			
ПКН-6 Способность анализировать и прогнозировать основные социально-экономические показатели, предлагать стратегические направления экономического развития на микро-, мезо- и макроуровнях	1. Применяет методический инструментарий системного анализа и моделирования экономических процессов для обоснования внедрения инновационных разработок с целью получения конкурентных преимуществ и обеспечения опережающего роста на новых и развивающихся рынках	Знать: методы решения задач Data Mining, применимых для анализа и прогноза основных социально-экономических показателей Уметь: применять методы машинного обучения для анализа и моделирования экономических процессов для обоснования внедрения инновационных разработок	Задание 1 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, проанализировать данные временных рядов с использованием библиотеки D-Tale. Задание 2 Решить задачу анализа временных рядов с использованием библиотеки Prophet
	2. Обосновывает перспективы изменений основных социально-экономических показателей и стратегические направления экономического развития на микро-, мезо- и макроуровнях	Знать: метрики качества и способы объяснения моделей машинного обучения, применяемых для анализа и прогнозирования основных социально-экономических показателей Уметь: обосновывать принимаемые решения, полученные на основе применения методов машинного обучения на различных уровнях	Задание 2 Провести интерпретацию результатов решения задачи анализа временных рядов с использованием библиотеки ARIMA Задание 2 Провести интерпретацию результатов решения задачи анализа временных рядов с использованием библиотеки Darts Задание 3 Решить задачу анализа временных рядов с использованием (на выбор) библиотек python: Auto_TS, sktime, TSFresh, Statsforecast; оценить модели, провести

			интерпретацию результатов
ПК-6 Способность применять методы интеллектуального анализа данных к решению прикладных задач, связанных с оказанием финансовых услуг	1. Владеет основными методами интеллектуального анализа данных и машинного обучения, используемые в экономике и финансах.	Знать: основные методы интеллектуального анализа данных и машинного обучения, используемые в экономике и финансах Уметь: применять основные методы интеллектуального анализа данных и машинного обучения в экономике и финансах	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 Задание 2 В Colab провести очистку, трансформацию данных, выполнить preprocessing данных, провести исследовательский анализ данных с использованием библиотек визуализации данных Matplotlib, Seaborn и Plotly express
	2. Владеет профессиональной терминологией в области больших данных и машинного обучения	Знать: основные теоретические положения в области машинного обучения и анализа данных Уметь: использовать открытую документацию по платформам, фреймворкам и библиотекам машинного обучения и анализа данных	Задание 1 В репозитории TensorFlow Hub найти модули для обработки структурированных (табличных данных), используя метод трансферного обучения дообучить модель используя фреймворк TensorFlow Задание 2 В репозитории TensorFlow Hub найти модули для обработки табличных данных, используя метод трансферного обучения переобучить модель используя фреймворк Keras
	3. Владеет современными пакетами прикладных программ, реализующими методы	Знать: спектр доступных библиотек, фреймворков и платформ машинного обучения и аналитики больших данных, применяемых в	Задание 1 В Colab сравнивать алгоритмы регрессии, градиентного бустинга, случайного леса, используемые для создания модели машинного

	обработки больших данных и машинного обучения при оказании финансовых услуг	экономике и финансах Уметь: применять пакеты прикладных программ, библиотеки, фреймворки и платформы машинного обучения и аналитики больших данных при оказании финансовых услуг	обучения, применяя фреймворк scikit-learn. Провести прогнозирование ответа лучшей модели на новых данных Задание 2 В Colab используя библиотеку Vaex провести исследовательский анализ больших объемов структурированных данных
	4. Применяет модели машинного обучения для решения практических задач в области предоставления финансовых услуг	Знать: задачи и методы Data Mining, методiku проведения кросс-индустриального стандарта Data Mining (CRISP DM) Уметь: использовать библиотеки python для решения задач машинного обучения, включая получение данных, их предобработку, создание и обучение модели, тюнинг модели, оценку и объяснение модели, деплой модели с использованием NumPy, Pandas, Scikit-learn, применять фреймворки автоматизированного машинного обучения.	Задание 1 В Colab обучить и сравнивать модели машинного обучения с использованием технологий AutoML фреймворка h2o-3 используя алгоритмы регрессии, градиентного бустинга, случайного леса или используя платформы. Провести прогнозирование ответа лучшей модели на новых данных, осуществить оценку и объяснение модели, деплой модели Задание 2 В Colab сравнивать алгоритмы регрессии, градиентного бустинга, случайного леса, используемые для создания модели машинного обучения, применяя фреймворк LightGBM. Провести прогнозирование ответа лучшей модели на новых данных, осуществить оценку и объяснение модели, деплой модели
	5. Обладает навыками анализа и обработки финансовой информации, необходимой для решения задач в области финансового анализа, кредитного	Знать: особенности различных доступных библиотек, фреймворков и платформ машинного обучения для решения задач в области финансового анализа, кредитного скоринга Уметь: применять различные	Задание 1 Применяя фреймворки машинного обучения scikit-learn, LightGBM, XGBoost, CatBoost решить задачу кредитного скоринга, сравнить результаты Задание 2 Применяя фреймворки автоматического машинного обучения h2o-3, AutoGluon, PyCaret решить

	скоринга экономического прогнозирования	библиотеки, фреймворки и платформы машинного обучения для решения задач в области финансового анализа, кредитного скоринга экономического прогнозирования	задачу в области финансового анализа, кредитного скоринга экономического прогнозирования, сравнить полученные результаты
«Анализ и стратегический менеджмент в бизнесе»			
ПК-2 Способность применять современные инструменты анализа, в том числе, с использованием интеллектуальных информационно-аналитических систем	1.Проводит самостоятельные исследования бизнес-процессов в соответствии с разработанной производственной программой с использованием современных информационно-аналитических систем.	Знать: основные библиотеки и фреймворки анализа данных и методы их использования в современных информационно-аналитических системах Уметь: применять основные библиотеки анализа данных в современных информационно-аналитических системах	Задание 1 В аналитической платформе Loginom, в узле Python, используя библиотеку pandas и sklearn, выполнить очистку и предобработку данных, результаты визуализировать встроенными средствами Loginom Задание 2 В платформе Power BI Desktop, в модуле Power Query, используя библиотеку pandas и sklearn, выполнить очистку и предобработку данных, результаты визуализировать средствами Power BI Desktop
	2.Демонстрирует владение методами сбора, анализа и обработки данных для принятия управленческих решений при моделировании бизнес-процессов	Знать: основные методы сбора, анализа и обработки данных с использованием библиотек python Уметь: применять основные методы сбора, анализа и обработки данных с использованием библиотек python для принятия управленческих решений	Задание 1 В Anaconda Jupyter Notebook, используя библиотеку python Bamboolib или Mitosheet (на выбор), провести исследовательский анализ данных финансового или экономического датасета, выявить скрытые закономерности с помощью создания модели машинного обучения Задание 2 В Colab используя библиотеку Pandas или Polars провести исследовательский анализ экономических (финансовых) данных

Примерные вопросы к зачету:

1. Охарактеризуйте отличия понятий искусственный интеллект, машинное обучение и глубокое обучение.
2. Охарактеризуйте процесс подготовки данных и разработки функций в машинном обучении.
3. Охарактеризуйте принципы работы машинного обучения: виды машинного обучения и минимизация функции потерь.
4. Охарактеризуйте метрики качества классификации: матрица ошибок (неточностей), Accuracy, Precision и Recall, F1-мера.
5. Охарактеризуйте основные возможности библиотеки Polar.
6. Охарактеризуйте ограничения библиотеки Polar.
7. Охарактеризуйте основные возможности библиотеки Vaex.
8. Охарактеризуйте ограничения библиотеки Vaex.
9. Охарактеризуйте основные возможности библиотеки Dask.
10. Охарактеризуйте ограничения библиотеки Dask.
11. Охарактеризуйте основные возможности библиотеки Modin.
12. Охарактеризуйте ограничения библиотеки Modin.
13. Охарактеризуйте основные возможности библиотеки Ray.
14. Охарактеризуйте ограничения библиотеки Ray.

Примерные практические задания к зачету:

1. Разработка сценариев ABC-XYZ, RFM, когортного - анализов в выбранной среде.
2. Визуализация результатов аналитики: в библиотеки Altair, Plotly, а также встроенными средствами платформ Deepnote (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 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: 18.12.2023). – 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: 18.12.2023). – 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: 18.12.2023). – 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: 18.12.2023). – 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>

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.