

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Андрей Драгомирович Хлутков  
Должность: директор  
Дата подписания: 16.06.2026 20:27:56  
Уникальный программный ключ:  
880f7c07c583b07b775f6604a630281b13ca9fd2

Приложение 4  
к образовательной программе

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

**Б1.В.ДЭ.01.01 «Методы анализа данных (продвинутый уровень)»**  
(индекс, наименование дисциплины в соответствии с учебным планом)

**38.04.05 «Бизнес-информатика»**  
(код, наименование направления подготовки/специальности)

**«Аналитическое обеспечение информационной безопасности»**  
\_\_\_\_\_  
(наименование образовательной программы)

очная  
(форма обучения)

Год набора 2026

Город  
Санкт-Петербург, 2026 г.

**Автор–составитель:**

Доктор военных наук профессор, заведующий кафедрой бизнес-информатики  
Наумов Владимир Николаевич

**Заведующий кафедрой бизнес-информатики**

Доктор военных наук, профессор Наумов Владимир Николаевич

РПД Б1.В.ДЭ.01.01 «Методы анализа данных (продвинутый уровень)»  
одобрена протоколом заседания кафедры бизнес-информатики № 6 от  
26.03.2026 г.

## СОДЕРЖАНИЕ

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы
2. Объем и место дисциплины в структуре образовательной программы
3. Содержание и структура дисциплины
4. Материалы текущего контроля успеваемости обучающихся
5. Оценочные материалы промежуточной аттестации по дисциплине
6. Методические материалы для освоения дисциплины

Учебная литература и ресурсы информационно-телекоммуникационной сети "Интернет"

7.1. Основная литература

7.2. Дополнительная литература

7.3. Нормативные правовые документы и иная правовая информация

7.4. Интернет-ресурсы

7.5. Иные источники

Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы

## 1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения программы

1.1. Дисциплина Б1.В.ДЭ.01.01 «Методы анализа данных (продвинутый уровень)» обеспечивает овладение следующими компетенциями.

Таблица 1.1

ОТФ/ТФ и реквизиты ПС (при наличии)**	Код компетенции **	Наименование Компетенции **	Код индикатора достижения компетенции **	Наименование индикатора достижения компетенции **	Образовательный результат **
<p>08.037 Бизнес-аналитик Е Управление бизнес-анализом</p> <p>ТД.3. Сбор информации, анализ, оценка эффективности проводимого бизнес-анализа в организации</p>	ПКс-3	Способен управлять работами по сопровождению и проектами по созданию (модификации) ИС, автоматизирующих задачи организационного управления и бизнес-процессы	ПКс-3.1	Управляет данными при решении задач бизнес-аналитики	<p>ПКс 3.1. <b>Знает</b></p> <ul style="list-style-type: none"> <li>– Возможности использования свободно распространяемого программного обеспечения в организации в объеме, необходимом для целей бизнес-анализа</li> <li>– Перспективные и существующие цифровые технологии и цифровые возможности для бизнеса в контексте предметной области и специфики деятельности организации</li> </ul> <p>ОПК-ПКс-3.1</p> <p><b>Умеет</b></p> <ul style="list-style-type: none"> <li>– Пользоваться системами анализа и визуализации данных;</li> <li>– Применять ИТ-инструменты (приложения и платформы) для обеспечения работ по бизнес-анализу</li> <li>– Отбирать, применять и адаптировать соответствующие методы,</li> </ul>

					инструменты и техники анализа бизнес-ситуации и предметной области, включая методы анализа данных
--	--	--	--	--	---

*\* Дисциплина может формировать компетенцию полностью или частично.*

*\*\*Должно соответствовать Приложению 1 к образовательной программе*

## **2.Объем и место дисциплины в структуре образовательной программы**

### **Объем дисциплины**

Общий объем дисциплины Б1.В.ДЭ.01.01 «Методы анализа данных (продвинутый уровень)» - 4 зачетные единицы – 144 акад.час; объем академических часов, выделенных на контактную работу обучающихся с преподавателем - 39 акад час, из них 12 акад.час – лекции, 16 час – практические занятия, 9 часов – контактная работа на аттестацию в период экзаменационных сессий, 2 акад.час – консультация к экзамену, 18 акад.час – самостоятельная работа при подготовке к экзамену и 87 акад. час. выделяется на самостоятельную работу обучающихся.

Дисциплина Б1.В.ДЭ.01.01 «Методы анализа данных (продвинутый уровень)» относится к элективным дисциплинам части образовательной программы формируемой участниками образовательных отношений, преподается на 1 курсе во 2 семестре, когда у обучаемыми получены знания и сформированы навыки, полученные на таких дисциплинах как:

Б1.В.09 «Интеллектуальный анализ текстов и изображений», Б1.О.07 «Аналитическая поддержка принятия решений», Б1.В.ДЭ.02.01 «Менеджмент данных» и является основой для подготовки к магистерским диссертациям.

### 3.Содержание и структура дисциплины

Таблица 3

#### 3.1. Структура дисциплины Очная форма обучения

№ п/п	Наименование тем и (или) разделов	ВСЕ- ГО	Объем дисциплины, ак.час										Форма текущего  контроля успеваемости, промежуточ- ной аттеста- ции	
			Контактная работа обучающихся  с преподавателем по видам учебных занятий							Самостоятель- ная работа				
			Период теоретического обучения					Период промежуточ- ной аттестации (сес- сия)						
			Занятия лекцион- ного типа		Занятия семинар- ского типа		И К	КС Р	К Э	Каттэ к	Контрол ь	СРк р		СРэ к
Л	ВЛ	ЛР	ПЗ											
Тема 1	Продвинутая	32	4			4							24	Т, ПКЗ

	подготовка данных и исследовательский анализ (EDA)														
Тема 2	Структурное моделирование (PLS-SEM)	30	4			4								22	ПКЗ, Т
Тема 3	Сетевой анализ (Network Analysis)	27	2			4								21	ПКЗ, Т
Тема 4	Генеративный ИИ и этика в бизнес-аналитике	26	2			4								20	Т
Промежуточная аттестация		29						2	9				18		Экзамен
Итого		144	12			16	0	0		4				87	

*Используемые сокращения:*

Л – лекции - занятия, предусматривающие преимущественную передачу учебной информации обучающимся педагогическими работниками организации и (или) лицами, привлекаемыми организацией к реализации образовательных программ на иных условиях,).

ВЛ – видео лекции.

ЛР – лабораторные работы.

ПЗ – практические занятия (за исключением лабораторных работ).

ИК – индивидуальные консультации.

КСР – контроль самостоятельной работы

КЭ – консультации перед экзаменом

Каттэк – контактная работа на аттестацию в период экзаменационных сессий

Контроль - контактная работа на аттестацию в период экзаменационных сессий для заочной формы обучения

СРкр – самостоятельная работа на подготовку курсовой работы/ курсового проекта.

СРэк – самостоятельная работа на подготовку к экзамену.

СР – самостоятельная работа в семестре на подготовку к учебным занятиям.

Т – тестирование.

ПКЗ – практические контрольные задания.

ПИЗ – профессионально-исследовательские задания.

### 3.2 Содержание дисциплины

#### Тема 1. Продвинутое подготовка данных и исследовательский анализ (EDA)

Типы пропусков: MCAR, MAR, MNAR — определения, примеры, последствия для анализа. Методы обработки пропусков: удаление (listwise/pairwise), простая импутация (mean/median/mode), множественное импутирование (MICE). Диагностика механизма пропусков: визуализация (paniar), тест Литтла, логистические модели предсказания пропусков. Методы обнаружения выбросов: статистические (IQR, z-score), модельные (Isolation Forest, Local Outlier Factor), визуальные (boxplot, scatterplot matrix). Бизнес-интерпретация аномалий: ошибка измерения vs редкое событие vs мошенничество. Особенности EDA для опросных данных: шкалы Лайкерта, обратное кодирование, проверка направленности индикаторов

- Анализ распределений: асимметрия, эксцесс, тесты нормальности (Shapiro-Wilk, Kolmogorov-Smirnov), графические методы (QQ-plot, histogram + density). Корреляционный анализ: Пирсон, Спирмен, полихорические корреляции для порядковых данных; визуализация (corrplot, GGally::ggpairs).

#### Тема 2. Структурное моделирование (PLS-SEM)

Ковариационный (CB-SEM) vs дисперсионный (PLS-SEM) подходы: цели, допущения, область применения. Рефлексивная и формативная спецификация индикаторов: теоретическое обоснование, эмпирические следствия. Построение концептуальной рамки: экзогенные/эндогенные конструкты, прямые/косвенные связи. Требования к выборке, мощности теста, соотношению наблюдений и индикаторов. Экосистема программного обеспечения: R (cSEM, semnr), SmartPLS, WarpPLS; выбор инструмента под задачу. Предварительная оценка факторной структуры: параллельный анализ (parallel analysis), скрин-тест, критерий Кайзера.

Проверка допущений для SEM: линейность, гомоскедастичность, отсутствие идеальной мультиколлинеарности (VIF), нормальность остатков. Надёжность шкал: факторные нагрузки (loadings), составная надёжность (CR), Cronbach's  $\alpha$ . Конвергентная валидность: средняя дисперсия (AVE), пороговые критерии. Дискриминантная валидность: кросс-нагрузки, критерий Fornell-Larcker, современный стандарт HTMT. Валидация формативных конструктов: значимость весов, релевантность, мультиколлинеарность (VIF), внешняя валидность. Стратегии корректировки модели: удаление/замена индикаторов, агрегация, теоретическое обоснование изменений

#### Тема 3. Сетевой анализ (Network Analysis)

Базовые понятия теории графов: узлы (вершины), рёбра (связи),

направленность, взвешенность, петли, мультиграфы. Типы сетей в бизнес-контексте: социальные (сотрудники, клиенты), транзакционные (платежи, поставки), информационные (цитирование, коммуникация), технологические (зависимости процессов).• Форматы представления сетей: матрица смежности, список рёбер (edge list), инцидентность; конвертация форматов в R (igraph, tidygraph). Сбор и подготовка сетевых данных: источники (логи коммуникаций, CRM, ERP), этические аспекты, анонимизация, работа с пропусками в связях.

Локальные метрики: степень (degree), взвешенная степень, полустепени (in/out-degree). Метрики посредничества: betweenness centrality — алгоритм, интерпретация, вычислительная сложность, нормализация. Метрики близости: closeness centrality, harmonic centrality — применение в анализе скорости распространения информации. Метрики влияния: eigenvector centrality, PageRank — логика рекурсивного влияния, применение для выявления ключевых агентов. Дополнительные метрики: коэффициент кластеризации, плотность, диаметр, эксцентриситет, k-core decomposition. Бизнес-интерпретация: выявление «узких мест», информационных брокеров, изолированных групп, уязвимостей в цепочках поставок

#### **Тема 4. Генеративный ИИ и этика в бизнес-аналитике**

Архитектуры генеративных моделей: трансформеры, диффузионные модели, LLM — принципы работы и ограничения.

Типы задач для бизнес-аналитики: генерация кода, синтез данных, автоматизация отчётности, нарративная аналитика, симуляция сценариев. Интеграция GenAI в аналитический цикл: EDA-ассистенты, автодокументирование, генерация гипотез, проверка кода (GitHub Copilot, CodeLlama, R-assistants). Оценка качества выводов ИИ: метрики релевантности, полноты, согласованности; выявление «галлюцинаций». Практические ограничения: стоимость, задержки, контекстное окно, воспроизводимость промптов.

#### **4. Типы оценочных материалов, показатели и критерии оценивания**

Оценочные материалы по дисциплине Б1.В.ДЭ.01.01 «Методы анализа данных (продвинутый уровень)» входят в состав оценочных материалов по образовательной программе. Совокупность оценочных материалов по всем дисциплинам образовательной программы составляет фонд оценочных средств (далее – ФОС). ФОС используется при проведении текущего контроля успеваемости и промежуточной аттестации обучающихся с целью оценивания достижения обучающимися планируемых результатов обучения.

4.2. ФОС разработан как комплекс проверочных заданий различного типа и уровня сложности, включает критерии и шкалы оценивания, а также

«ключи» правильных ответов. ФОС формируется как отдельный документ и хранится в электронном виде, доступ к ФОС предоставлен ограниченному кругу лиц.

4.3. Для самостоятельной работы обучающихся при подготовке к текущему контролю успеваемости и промежуточной аттестации в рабочих программах дисциплин размещены типовые проверочные задания, которые можно условно разделить на задания закрытого, комбинированного и открытого типов.

Задания закрытого типа — это тестовые задания, в которых каждый вопрос сопровождается готовыми вариантами ответов, из которых необходимо выбрать один или несколько правильных.

Задания комбинированного типа – это тестовые задания, в которых каждый вопрос сопровождается готовыми вариантами ответов, из которых необходимо выбрать один или несколько правильных и обосновать свой выбор.

Задания открытого типа — это задания, в которых на каждый вопрос должен быть предложен развернутый обоснованный ответ.

В зависимости от типа задания рекомендованы определенная последовательность выполнения и система оценивания выполнения заданий.

#### 4.4. Типы заданий, сценарии выполнения, критерии оценивания

ТИП ЗАДАНИЯ	ИНСТРУКЦИЯ	СЦЕНАРИИ ВЫПОЛНЕНИЯ	КРИТЕРИИ ОЦЕНИВАНИЯ
Задание закрытого типа с выбором одного правильного ответа из нескольких вариантов предложенных	Прочитайте текст, выберите правильный ответ	<ol style="list-style-type: none"> <li>1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается только один из предложенных вариантов.</li> <li>2. Внимательно прочитать предложенные варианты ответа.</li> <li>3. Выбрать один верный ответ.</li> <li>4. Записать только номер (или букву) выбранного варианта ответа (например, 3 или В).</li> </ol>	Ответ считается верным, если правильно указана цифра или буква
Задание закрытого типа на установление соответствия	Прочитайте текст и установите соответствие	<ol style="list-style-type: none"> <li>1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидаются пары элементов.</li> <li>2. Внимательно прочитать оба списка: список 1 – вопросы, утверждения, факты, понятия и т.д.; список 2 – утверждения, свойства объектов и т.д.</li> <li>3. Сопоставить элементы списка 1 с элементами списка 2, сформировать пары элементов.</li> <li>4. Записать попарно буквы и цифры (в зависимости от задания) вариантов ответа (например, А1 или Б4).</li> </ol>	Ответ считается верным, если правильно указаны цифры или буквы
Задание закрытого типа с выбором нескольких правильных ответов из нескольких вариантов предложенных	Прочитайте текст, выберите правильные ответы	<ol style="list-style-type: none"> <li>1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается несколько правильных ответов из предложенных вариантов.</li> <li>2. Внимательно прочитать предложенные варианты ответа.</li> <li>3. Выбрать несколько правильных ответов.</li> <li>4. Записать только номера (или буквы) выбранного варианта ответа (например, 1 4 или А Г).</li> </ol>	Ответ считается верным, если правильно установлены все соответствия (позиции из одного столбца верно сопоставлены с позициями другого)
Задание закрытого типа на установление последовательности	Прочитайте текст и установите последовательность	<ol style="list-style-type: none"> <li>1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается последовательность элементов.</li> </ol>	Ответ считается верным, если правильно указана вся последовательность цифр

		<p>2. Внимательно прочитать предложенные варианты ответа.</p> <p>3. Построить верную последовательность из предложенных элементов.</p> <p>4. Записать буквы/цифры (в зависимости от задания) вариантов ответа в нужной последовательности (например, БВА или 135).</p>	
<p>Задание комбинированного типа с выбором одного правильного ответа из предложенных и обоснованием выбора</p>	<p>Прочитайте текст, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа</p>	<p>1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается только один из предложенных вариантов.</p> <p>2. Внимательно прочитать предложенные варианты ответа.</p> <p>3. Выбрать один верный ответ.</p> <p>4. Записать только номер (или букву) выбранного варианта ответа.</p> <p>5. Записать аргументы, обосновывающие выбор ответа (например, 4 текст обоснования).</p>	<p>Ответ считается верным, если правильно указана цифра или буква и приведены корректные аргументы, используемые при выборе ответа</p>
<p>Задание открытого типа с развернутым ответом</p>	<p>Прочитайте текст и запишите развернутый обоснованный ответ</p>	<p>1. Внимательно прочитать текст задания и понять суть вопроса.</p> <p>2. Продумать логику и полноту ответа.</p> <p>3. Записать ответ, используя четкие компактные формулировки.</p> <p>4. В случае расчетной задачи, записать решение и ответ</p>	<p>Ответ считается верным:</p> <p>1. Отсутствие фактических ошибок.</p> <p>2. Раскрытие объема используемых понятий (полнота ответа).</p> <p>3. Обоснованность ответа (наличие аргументов).</p> <p>4. Логическая последовательность излагаемого материала.</p>

#### 4.5. Общая шкала оценивания результатов текущего контроля успеваемости и промежуточной аттестации обучающихся с применением БРС

Итоговая балльная оценка	Традиционная система	Бинарная система	ECTS	
			Для традиционной системы	Для бинарной системы
95-100	Отлично	Зачтено	A	P/ Passed
85-94			B	P/ Passed
75-84	Хорошо		C	P/ Passed
65-74			D	P/ Passed
55-64	Удовлетворительно		E	P/ Passed
0-54	Неудовлетворительно	Не зачтено	F	F/Failed

Соотношение баллов за текущий контроль успеваемости и промежуточную аттестацию, а также повторную промежуточную аттестацию:

Максимальная сумма баллов за текущий контроль успеваемости	Максимальная сумма баллов за промежуточную аттестацию	Максимальная итоговая балльная оценка	Максимальная сумма баллов за повторную промежуточную аттестацию
60 баллов	40 баллов	100 баллов	100баллов

### 3. Формы аттестации, типовые оценочные материалы для текущего контроля успеваемости обучающихся, критерии и шкалы оценивания по контрольным точкам

5.1. В ходе реализации дисциплины используются следующие формы текущего контроля успеваемости обучающихся (в том числе, задания к контрольным точкам):

T – тестирование, ПКЗ – практические контрольные задания.

#### Тема 1. Продвинутая подготовка данных и исследовательский анализ (EDA)

Тестовые задания:

*Задание закрытого типа с выбором одного правильного ответа из нескольких вариантов предложенных.*

1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается только один из предложенных вариантов.
2. Внимательно прочитать предложенные вариант-ты ответа.
3. Выбрать один верный ответ.

4. Записать только номер (или букву) выбранного варианта ответа (например, 3 или B)..

**Тест 1.** Какой механизм пропусков является **НАИБОЛЕЕ** проблематичным для анализа и требует специализированных методов?

A) MCAR (Missing Completely At Random)

B) MAR (Missing At Random)

C) MNAR (Missing Not At Random) — вероятность пропуска зависит от самого ненаблюдаемого значения

D) Все механизмы одинаково просты в обработке

Правильный ответ: C

Обоснование: При MNAR пропуск систематически связан с ненаблюдаемым значением (напр., люди с высоким доходом реже указывают зарплату). Это ведёт к смещению оценок; требуются selection models, sensitivity analysis или экспертные допущения.

**Тест 2.** В чём основное преимущество множественного импутирования (MICE) перед простым удалением строк с пропусками?

A) Оно быстрее выполняется

B) Оно сохраняет размер выборки и учитывает неопределённость импутации через создание нескольких датасетов

C) Оно автоматически исправляет все выбросы

D) Оно не требует проверки допущений

Правильный ответ: B

Обоснование: MICE генерирует  $m$  импутированных датасетов, анализирует каждый и объединяет результаты (Rubin's rules). Это сохраняет статистическую мощь и корректно оценивает стандартные ошибки, в отличие от listwise deletion.

**Тест 3.** Какой метод обнаружения выбросов наиболее устойчив к многомерным аномалиям?

A) Правило  $1.5 \times IQR$  для одномерных боксплотов

B) Z-score  $> |3|$  для нормальных распределений

C) Isolation Forest — алгоритм, изолирующий аномалии через случайные разбиения пространства

D) Визуальный осмотр рассеянных диаграмм

Правильный ответ: C

Обоснование: Isolation Forest эффективно выявляет многомерные выбросы, не требуя предположений о распределении. Он основан на том, что аномалии «изолируются» быстрее при случайных разбиениях, чем нормальные наблюдения.

Практическое контрольное задание по теме 1

Тема: «Комплексный EDA датасета опроса о восприятии ИИ»

Загрузить датасет (предоставляется: ~500 наблюдений, 25 переменных, включая пропуски).

Выполнить разведочный анализ:

Описать структуру данных, типы переменных, долю пропусков.

Визуализировать распределения ключевых переменных (гистограммы, boxplot).

Построить корреляционную матрицу с выделением значимых связей.

Применить `mi` для множественного импутирования ( $m=5$ ), сравнить результаты с полным удалением пропусков.

Выявить и обосновать обработку выбросов (мин. 2 метода).

Подготовить отчёт в R Markdown с интерпретацией выводов для менеджера.

## Тема 2. Структурное моделирование (PLS-SEM)

### Тестовые задания:

*Задание закрытого типа с выбором одного правильного ответа из нескольких вариантов предложенных.*

1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается только один из предложенных вариантов.

2. Внимательно прочитать предложенные варианты-ты ответа.

3. Выбрать один верный ответ.

4. Записать только номер (или букву) выбранного варианта ответа (например, 3 или B).

*Задание закрытого типа с выбором одного правильного ответа из нескольких вариантов предложенных.*

1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается только один из предложенных вариантов.

2. Внимательно прочитать предложенные варианты-ты ответа.

3. Выбрать один верный ответ.

4. Записать только номер (или букву) выбранного варианта ответа (например, 3 или B).

**Тест 1.** В чём ключевое методологическое различие между ковариационным (CB-SEM) и дисперсионным (PLS-SEM) подходами?

A) CB-SEM минимизирует остатки дисперсии, PLS-SEM — ковариации

B) CB-SEM ориентирован на подтверждение теории и требует нормальности данных; PLS-SEM максимизирует объяснённую дисперсию, устойчив к отклонениям от нормальности и работает с малыми выборками;

C) PLS-SEM требует строгой многомерной нормальности, CB-SEM — нет

D) Различий нет, это синонимы одного алгоритма

Правильный ответ: B

Обоснование: CB-SEM (covariance-based) оценивает, насколько модель воспроизводит эмпирическую ковариационную матрицу, требует больших  $N$  и нормальности. PLS-SEM (variance-based) последовательно оценивает конструкты, ориентирован на прогнозирование, менее требователен к распределению и размеру выборки.

**Тест 2.** Какой показатель является стандартом оценки конвергентной

валидности рефлексивного конструкта?\*\*\*

- A) Cronbach's  $\alpha \geq 0.7$
- B) Average Variance Extracted (AVE)  $\geq 0.50$
- C) HTMT  $< 0.85$
- D) SRMR  $< 0.08$

Правильный ответ: B

Обоснование: AVE показывает долю дисперсии индикаторов, объясняемую конструктом. Значение  $\geq 0.50$  означает, что конструкт объясняет  $\geq 50\%$  вариации своих показателей, что подтверждает конвергентную валидность.

**Тест 3.** Когда конструкт следует специфицировать как формативный, а не рефлексивный?

- A) Когда индикаторы коррелируют между собой  $> 0.8$
- B) Когда индикаторы совместно формируют конструкт, и удаление одного меняет его смысловое содержание
- C) Когда все факторные нагрузки  $> 0.7$
- D) Когда выборка  $< 100$  наблюдений

Правильный ответ: B

Обоснование: В формативной спецификации индикаторы \*вызывают\* латентную переменную (напр., «цифровая зрелость» = ИТ-инфраструктура + навыки + процессы). Удаление индикатора меняет суть конструкта. В рефлексивной — конструкт вызывает индикаторы.

**Практическое Контрольное задание 2.** Тема: «Построение и валидация модели влияния когнитивной нагрузки на качество решений при использовании ИИ»

На основе теоретической рамки (предоставляется) специфицировать модель в cSEM/semPLS: 3 экзогенных конструкта (доверие к ИИ, цифровая грамотность, сложность задачи); 1 эндогенный конструкт (качество решений)

Гипотеза: когнитивная нагрузка медиатирует влияние доверия на качество решений

Оценить измерительную модель: loadings, CR, AVE, HTMT.

Оценить структурную модель: пути,  $R^2$ ,  $f^2$ ,  $Q^2$ , SRMR.

Протестировать медиацию через бутстрэппинг (5000 итераций).

Подготовить отчёт с визуализацией модели (semPlot) и бизнес-интерпретацией.

### **Тема 3. Сетевой анализ (Network Analysis)**

#### **Тест по теме 3**

*Задание закрытого типа с выбором одного правильного ответа из нескольких вариантов предложенных.*

1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается только один из предложенных вариантов.
2. Внимательно прочитать предложенные варианты-ты ответа.
3. Выбрать один верный ответ.
4. Записать только номер (или букву) выбранного варианта ответа (например, 3 или B)..

**Тест 1.** Что представляет собой «ребро» (edge) в контексте сетевого анализа бизнес-процессов?\*

- A) Отдельный сотрудник или департамент в организационной структуре
- B) Связь или взаимодействие между двумя узлами сети (напр., коммуникация, транзакция)

C) Метрика центральности, измеряющая влияние узла

D) Визуальный элемент графа, не несущий аналитической нагрузки

Правильный ответ: B

Обоснование: Ребро (edge) формализует отношение между узлами: в бизнес-контексте это может быть передача задачи, финансовая транзакция, совместный проект. Именно рёбра позволяют анализировать потоки информации, ресурсов и влияния.

**Тест 2.** В каком случае для моделирования бизнес-сети следует использовать ОРИЕНТИРОВАННЫЙ граф?

A) Когда анализируются симметричные отношения (напр., «коллеги по отделу»)

B) Когда направление связи имеет значение (напр., «подчиняется», «отправляет отчёт», «платит»)

C) Когда сеть содержит менее 50 узлов

D) Когда все рёбра имеют одинаковый вес

Правильный ответ: B

Обоснование: Ориентированные графы (directed graphs) необходимы, когда связь асимметрична: влияние, поток ресурсов или информации имеют направление. Это критично для анализа иерархий, цепочек поставок, коммуникационных паттернов.

**Тест 3.** Какой формат представления сетевых данных НАИБОЛЕЕ эффективен для разреженных сетей (мало связей относительно числа узлов)?

A) Матрица смежности (adjacency matrix)

B) Список рёбер (edge list)

C) Полная таблица всех возможных пар узлов

D) Текстовое описание связей в свободной форме

Правильный ответ: B

Обоснование: Список рёбер хранит только существующие связи, что экономит память и ускоряет обработку для разреженных сетей (типично для бизнес-данных). Матрица смежности требует  $O(n^2)$  памяти независимо от числа рёбер.

**Тест 4.** Что означает «взвешенное ребро» в бизнес-сетевом анализе?

A) Ребро, помеченное цветом для визуализации

B) Ребро, которому присвоено числовое значение, отражающее интенсивность/важность связи (напр., частота коммуникаций, объём транзакций)

C) Ребро, соединяющее узлы из разных сообществ

D) Ребро, добавленное искусственно для заполнения пропусков

Правильный ответ: В

Обоснование: Вес ребра количественно характеризует силу связи. В бизнес-аналитике это позволяет ранжировать взаимодействия, выявлять ключевые каналы влияния, учитывать интенсивность потоков при расчёте метрик.

### **Практическое контрольное задание по теме 3**

Тема: «Анализ сети коллабораций в инновационном проекте»

Загрузить данные о взаимодействиях в команде (матрица смежности или edge list).

Построить граф в igraph, визуализировать с учётом атрибутов узлов (роль, департамент).

Расчитать метрики центральности для всех узлов, выделить топ-5 ключевых агентов.

Обнаружить сообщества (Louvain), интерпретировать их с учётом бизнес-контекста.

Оценить плотность, диаметр, коэффициент кластеризации сети; сформулировать рекомендации по оптимизации коммуникаций.

### **Тема 4. Генеративный ИИ и этика в бизнес-аналитике**

**Тест 1.** Какая архитектура лежит в основе большинства современных больших языковых моделей (LLM), используемых для бизнес-аналитики?

- A) Сверточные нейронные сети (CNN)
- B) Рекуррентные нейронные сети (RNN)
- C) Трансформеры с механизмом внимания (Transformer with attention)
- D) Генеративно-состязательные сети (GAN)

Правильный ответ: С

Обоснование: Архитектура Transformer с механизмом self-attention обеспечивает эффективную обработку длинных последовательностей и контекстуальное понимание, что критично для аналитических задач. CNN/RNN уступают в масштабируемости, GAN ориентированы на генерацию изображений.

**Тест 2.** Какой из перечисленных сценариев НАИМЕНЕЕ подходит для применения генеративного ИИ в бизнес-аналитике?

- A) Автоматическая генерация кода для визуализации данных в R
- B) Синтез реалистичных тестовых данных для отладки моделей
- C) Принятие финального решения о кредитовании клиента на основе сгенерированного отчёта
- D) Формулировка гипотез для дальнейшего статистического тестирования

Правильный ответ: С

Обоснование: Генеративный ИИ не должен принимать автономные решения в высокорисковых сценариях (кредитование, найм, медицина) из-за риска галлюцинаций, смещений и отсутствия юридической ответственности. ИИ выступает как ассистент, а не как финальный арбитр.

**Тест 3. Что означает термин «галлюцинация» в контексте генеративного ИИ?**

- A) Ошибка в вычислениях из-за численной нестабильности
- B) Генерация правдоподобного, но фактически неверного или выдуманного ответа
- C) Переобучение модели на тренировочных данных
- D) Отказ модели отвечать на запрос из-за этических ограничений

Правильный ответ: B

Обоснование: Галлюцинация — это специфический артефакт LLM, когда модель генерирует уверенный, но недостоверный ответ. Это требует обязательной верификации выводов через источники или экспертную проверку.

**Практическое Контрольное задание по теме 4**

Тема: «Разработка этического гайдлайна для внедрения ИИ-аналитики в компании»

Задание:

Выбрать бизнес-сценарий (например, скоринг клиентов, подбор персонала, прогноз спроса).

С использованием GenAI (с обязательной верификацией) подготовить черновик гайдлайна, включающий:

Описание рисков (смещение, приватность, интерпретируемость)

Протоколы валидации модели и мониторинга дрейфа

Правила коммуникации выводов стейкхолдерам

Механизмы эскалации и человеческого контроля

Критически отредактировать черновик: выявить и исправить потенциальные галлюцинации, добавить ссылки на регуляторные источники.

Представить гайдлайн в формате интерактивного отчёта (Quarto) с возможностью экспорта в PDF.

5.2. Типовые оценочные материалы для текущего контроля успеваемости обучающихся (вне контрольных точек) приведены в п.6.2.

5.3. Один или несколько тематических блоков дисциплины завершаются контрольной точкой (далее – КТ). Текущий контроль успеваемости по дисциплине предусматривает не менее 2 (двух) и не более 10 (десяти) КТ в течение периода освоения дисциплины.

Максимальное количество баллов за любой тип работ в рамках КТ составляет 100 (сто) баллов.

Распределение весовых коэффициентов по КТ в рамках текущего контроля успеваемости по дисциплине и формулы расчета:

Наименование контрольной точки	Максимальное количество баллов за работу в рамках КТ,	Коэффициент веса контрольной точки	Результат контрольной точки, участвующий в формировании итоговой
--------------------------------	---	------------------------------------	--

	которое может набрать студент		балльной оценки по дисциплине (отражается в журнале БРС в СДО)
КТ - 1	100	0,15	15
КТ - 2	100	0,15	15
КТ-3	100	0,15	15
КЕ-4	100	0,15	15
Итого:	x	0,6	60

Формула расчета результата контрольной точки:

Результат контрольной точки = Количество баллов за работу в рамках КТ X  
Коэффициент веса контрольной точки.

5.4 Формы текущего контроля успеваемости обучающихся в рамках КТ и типовые оценочные материалы:

#### **КТ-1**

##### **Тема 1.**

Тестирование.

Практическое-контрольное задание (ПКЗ).

##### **Тема 2.**

Тестирование.

Практическое-контрольное задание (ПКЗ).

##### **Тема 3**

Тестирование.

Практическое-контрольное задание (ПКЗ).

##### **Тема 4.**

Тестирование

Практическое-контрольное задание (ПКЗ).

Для каждой формы текущего контроля успеваемости обучающихся в рамках КТ определены критерии оценивания результатов выполнения задания.

#### *1. Критерии оценивания тестирования за каждый тест в КТ1:*

Критерии оценки	Диапазон баллов	Описание критерия
<i>Количество правильных ответов</i>	<i>0</i>	<i>Количество правильных ответов менее 55%</i>
	<i>25</i>	<i>Количество правильных ответов от 55% до 64%</i>
	<i>50</i>	<i>Количество правильных ответов от 65% до 74%</i>
	<i>75</i>	<i>Количество правильных ответов от</i>

		75% до 84%
	100	Количество правильных ответов от 85% до 100%
Итого максимально:	100	

## 2. Критерии оценивания ПКЗ:

Критерии оценки	Диапазон баллов	Описание критерия
<i>Содержание и раскрытие выбранных понятий</i>	41-70	<i>Детальное, последовательное описание всех понятий на примере выбранной системы</i>
	21-40	<i>Поверхностное описание без привязки к выбранной системе</i>
	0-20	<i>Понятия раскрыты минимально или не раскрыты вовсе</i>
<i>Количество выполненных заданий</i>	30	<i>Количество выполненных заданий от 85% до 100%</i>
	15	<i>Количество выполненных заданий от 55% до 84%</i>
	0	<i>Количество выполненных заданий менее 55%</i>
Итого максимально:	100	

## 6. Формы промежуточной аттестации, критерии и шкала оценивания, типовые оценочные материалы по дисциплине

### 6.1. Промежуточная аттестация проводится в форме экзамена

Зачет проводится путем выполнения итогового теста в ЭО/ДОТ, где могут быть использованы следующие формы: устно в ДОТ - в форме обоснованных ответов на задания различного типа; письменно в СДО - в форме письменного решения заданий различного типа; тестирование в СДО.

### 6.2. Типовые оценочные материалы промежуточной аттестации.

Типовые оценочные материалы промежуточной аттестации

Вопросы к экзамену по дисциплине «Методы анализа данных (продвинутый уровень)»

1. Классификация методов анализа данных по целям (описательный, диагностический, прогнозный, пресскриптивный). Приведите бизнес-примеры.

2. В чём различие между подтвердительным (confirmatory) и разведочным (exploratory) анализом? Когда какой подход уместен?

3. Какие допущения лежат в основе структурного моделирования и как проверять их нарушение?

4. Что такое репродуцируемость в аналитике и какие инструменты (Git, R Markdown, Docker) обеспечивают её на практике?
5. Сравните подходы CB-SEM и PLS-SEM: цели, требования к данным, интерпретация результатов.
6. Как специфицировать формативный конструкт в измерительной модели? Какие риски возникают при ошибочной спецификации?
7. Опишите процедуру оценки дискриминантной валидности через НТМТ. Почему она предпочтительнее критерия Форнелла-Ларкера?
8. В каких бизнес-задачах сетевой анализ даёт инсайты, недоступные традиционной регрессии? Приведите пример.
9. Как интерпретировать метрику betweenness centrality для выявления «узких мест» в организационной коммуникации?
10. Сравните подходы CB-SEM и PLS-SEM: цели, требования к данным, интерпретация результатов.
11. Как специфицировать формативный конструкт в измерительной модели? Какие риски возникают при ошибочной спецификации?
12. Опишите процедуру оценки дискриминантной валидности через НТМТ. Почему она предпочтительнее критерия Форнелла-Ларкера?
13. В каких бизнес-задачах сетевой анализ даёт инсайты, недоступные традиционной регрессии? Приведите пример.
14. Как интерпретировать метрику betweenness centrality для выявления «узких мест» в организационной коммуникации?
15. Какие этапы аналитического пайплайна можно эффективно усилить с помощью генеративного ИИ? Где его применение рискованно?
16. Опишите практический протокол верификации выводов, сгенерированных ИИ-ассистентом.
17. Какие метрики справедливости (fairness) вы бы выбрали для оценки модели кредитного скоринга? Обоснуйте выбор.
18. Как адаптировать презентацию результатов структурного моделирования для аудитории менеджеров, не знакомых с статистикой?
19. Какие требования к документации моделей предъявляют современные регуляторные рамки (AI Act, GDPR)?

Типовые проверочные задания для самоподготовки обучающегося к промежуточной аттестации:

ТИП ЗАДАНИЯ	СЦЕНАРИИ ВЫПОЛНЕНИЯ	ТИПОВЫЕ ЗАДАНИЯ
Задание закрытого типа с выбором одного правильного ответа из нескольких предложенных	<ol style="list-style-type: none"> <li>1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается только один из предложенных вариантов.</li> <li>2. Внимательно прочитать предложенные варианты ответа.</li> <li>3. Выбрать один верный ответ.</li> <li>4. Записать только номер (или</li> </ol>	<ol style="list-style-type: none"> <li>1. Что можно оценить с помощью информационного критерия Акайке?.               <ol style="list-style-type: none"> <li>a) стациональность;</li> <li>b) нестационарность;</li> <li>c) качество модели;</li> <li>d) наличие аномалий;</li> <li>e) критерий не применяется для анализа временных рядов;</li> </ol> </li> </ol>

	<p>букву) выбранного варианта ответа (например, 3 или В).</p>	<p>f) ошибки аппроксимации модели; g) критерий применяется только совместно с критерием Байеса.</p> <p>2 Модель ARIMA имеет вид (1,1,3). Выберите правильный ответ</p> <p>a). Проинтегрированная модель авторегрессии первого порядка, скользящего среднего третьего порядка.</p> <p>b). Модель авторегрессии третьего порядка, скользящего среднего первого порядка.</p> <p>c). Полиномиальная модель в остатках.</p> <p>d). Сезонная авторегрессионная модель в остатках.</p> <p>3. Как ведет себя доверительный интервал прогноза?</p> <p>a) Уменьшается при увеличении доверительной вероятности;</p> <p>b) Увеличивается при увеличении доверительной вероятности;</p> <p>c) Ширина доверительного интервала постоянная независимо от используемого метода;</p> <p>d) При решении задачи прогнозирования доверительный интервал не строится.</p>																						
<p>Задание закрытого типа на установление соответствия</p>	<p>1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидаются пары элементов.</p> <p>2. Внимательно прочитать оба списка: список 1 – вопросы, утверждения, факты, понятия и т.д.; список 2 – утверждения, свойства объектов и т.д.</p> <p>3. Сопоставить элементы списка 1 с элементами списка 2, сформировать пары элементов.</p> <p>4. Записать попарно буквы и цифры (в зависимости от задания) вариантов ответа (например, А1 или Б4).</p>	<p>1. Установите соответствие между вопросом и вариантом ответа</p> <table border="1" data-bbox="890 891 1465 1709"> <thead> <tr> <th>Вопрос</th> <th>Вариант ответа</th> </tr> </thead> <tbody> <tr> <td><math>R^2 = 0.32</math> для конструктора «Качество решений»</td> <td>А) Модель имеет предиктивную силу: предсказывает лучше простого среднего</td> </tr> <tr> <td><math>f^2 = 0.18</math> для пути «Доверие к ИИ → Качество решений»</td> <td>Б) Модель объясняет 32% дисперсии эндогенного конструктора — умеренная объясняющая сила</td> </tr> <tr> <td><math>Q^2 = 0.12</math> (через blindfolding)</td> <td>В) Предиктор имеет средний практический эффект на целевую переменную</td> </tr> <tr> <td>Бутстрэпированный 95% ДИ для <math>\beta</math>: [0.15; 0.42]</td> <td>Г) Эффект статистически значим: доверительный интервал не содержит ноль</td> </tr> <tr> <td>SRMR = 0.09</td> <td>Д) Приближённое соответствие модели данным находится на границе приемлемого</td> </tr> <tr> <td></td> <td>Е) Модель переобучена: высокая объясняющая способность при низкой прогнозной</td> </tr> </tbody> </table> <p>2. Установите соответствие между метрикой и критерием оценки качества модели</p> <table border="1" data-bbox="890 1771 1465 2080"> <thead> <tr> <th>Метрика</th> <th>Критерий оценки качества модели</th> </tr> </thead> <tbody> <tr> <td>Average Variance Extracted (AVE)</td> <td>А) <math>&lt; 0.85-0.90</math> — подтверждение дискриминантной валидности</td> </tr> <tr> <td>Composite Reliability (CR)</td> <td>В) <math>\geq 0.50</math> — конструктор объясняет <math>\geq 50\%</math> дисперсии индикаторов</td> </tr> <tr> <td>HTMT (Heterotrait-</td> <td>С) <math>\geq 0.7</math> — внутренняя согласованность шкалы</td> </tr> </tbody> </table>	Вопрос	Вариант ответа	$R^2 = 0.32$ для конструктора «Качество решений»	А) Модель имеет предиктивную силу: предсказывает лучше простого среднего	$f^2 = 0.18$ для пути «Доверие к ИИ → Качество решений»	Б) Модель объясняет 32% дисперсии эндогенного конструктора — умеренная объясняющая сила	$Q^2 = 0.12$ (через blindfolding)	В) Предиктор имеет средний практический эффект на целевую переменную	Бутстрэпированный 95% ДИ для $\beta$ : [0.15; 0.42]	Г) Эффект статистически значим: доверительный интервал не содержит ноль	SRMR = 0.09	Д) Приближённое соответствие модели данным находится на границе приемлемого		Е) Модель переобучена: высокая объясняющая способность при низкой прогнозной	Метрика	Критерий оценки качества модели	Average Variance Extracted (AVE)	А) $< 0.85-0.90$ — подтверждение дискриминантной валидности	Composite Reliability (CR)	В) $\geq 0.50$ — конструктор объясняет $\geq 50\%$ дисперсии индикаторов	HTMT (Heterotrait-	С) $\geq 0.7$ — внутренняя согласованность шкалы
Вопрос	Вариант ответа																							
$R^2 = 0.32$ для конструктора «Качество решений»	А) Модель имеет предиктивную силу: предсказывает лучше простого среднего																							
$f^2 = 0.18$ для пути «Доверие к ИИ → Качество решений»	Б) Модель объясняет 32% дисперсии эндогенного конструктора — умеренная объясняющая сила																							
$Q^2 = 0.12$ (через blindfolding)	В) Предиктор имеет средний практический эффект на целевую переменную																							
Бутстрэпированный 95% ДИ для $\beta$ : [0.15; 0.42]	Г) Эффект статистически значим: доверительный интервал не содержит ноль																							
SRMR = 0.09	Д) Приближённое соответствие модели данным находится на границе приемлемого																							
	Е) Модель переобучена: высокая объясняющая способность при низкой прогнозной																							
Метрика	Критерий оценки качества модели																							
Average Variance Extracted (AVE)	А) $< 0.85-0.90$ — подтверждение дискриминантной валидности																							
Composite Reliability (CR)	В) $\geq 0.50$ — конструктор объясняет $\geq 50\%$ дисперсии индикаторов																							
HTMT (Heterotrait-	С) $\geq 0.7$ — внутренняя согласованность шкалы																							

		<table border="1"> <tr> <td>Monotrait ratio)</td> <td></td> </tr> <tr> <td>Факторная нагрузка (loading) рефлексивного индикатора</td> <td><math>D) \geq 0.707</math> — индикатор сильно связан с конструктом (допустимо <math>\geq 0.6</math> на exploratory stage)</td> </tr> <tr> <td>VIF для формативных индикаторов</td> <td><math>E) &lt; 3-5</math> — отсутствие проблемной мультиколлинеарности</td> </tr> <tr> <td></td> <td><math>F) &gt; 0.08</math> — приемлемое соответствие модели (не относится к измерительной модели)</td> </tr> </table>	Monotrait ratio)		Факторная нагрузка (loading) рефлексивного индикатора	$D) \geq 0.707$ — индикатор сильно связан с конструктом (допустимо $\geq 0.6$ на exploratory stage)	VIF для формативных индикаторов	$E) < 3-5$ — отсутствие проблемной мультиколлинеарности		$F) > 0.08$ — приемлемое соответствие модели (не относится к измерительной модели)						
Monotrait ratio)																
Факторная нагрузка (loading) рефлексивного индикатора	$D) \geq 0.707$ — индикатор сильно связан с конструктом (допустимо $\geq 0.6$ на exploratory stage)															
VIF для формативных индикаторов	$E) < 3-5$ — отсутствие проблемной мультиколлинеарности															
	$F) > 0.08$ — приемлемое соответствие модели (не относится к измерительной модели)															
Задание закрытого типа с выбором нескольких правильных ответов из нескольких вариантов предложенных	<p>1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается несколько правильных ответов из предложенных вариантов.</p> <p>2. Внимательно прочитать предложенные варианты ответа.</p> <p>3. Выбрать несколько правильных ответов.</p> <p>4. Записать только номера (или буквы) выбранного варианта ответа (например, 1 4 или А Г).</p>	<p>Какие показатели используются для оценки конвергентной валидности рефлексивного конструкта в PLS-SEM?</p> <p>(Выберите 2–3 правильных ответа)</p> <p>А) Критерий Фурнелла-Ларкера</p> <p>В) Average Variance Extracted (AVE) <math>\geq 0.50</math></p> <p>С) Факторные нагрузки (loadings) <math>\geq 0.707</math></p> <p>Д) Индекс SRMR <math>&lt; 0.08</math></p> <p>Е) Составная надёжность (Composite Reliability, CR) <math>\geq 0.7</math></p> <p>Какие этические и методологические риски связаны с использованием генеративного ИИ при построении PLS-моделей?</p> <p>(Выберите 2–3 правильных ответа)</p> <p>А) Риск галлюцинаций: ИИ может сгенерировать неверный код для бутстрэппинга или интерпретации метрик</p> <p>В) Отсутствие воспроизводимости: стохастичность генерации затрудняет точное повторение анализа</p> <p>С) Автоматическое повышение точности модели за счёт ИИ-оптимизации гиперпараметров</p> <p>Д) Смещение в интерпретации: ИИ может предлагать причинные выводы на основе корреляционных данных</p> <p>Е) Полное исключение необходимости экспертной проверки при использовании сертифицированных ИИ-инструментов</p>														
Задание закрытого типа на установление последовательности	<p>1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается последовательность элементов.</p> <p>2. Внимательно прочитать предложенные варианты ответа.</p> <p>3. Построить верную последовательность из предложенных элементов.</p> <p>4. Записать буквы/цифры (в зависимости от задания) вариантов ответа в нужной последовательности (например, БВА или 135).</p>	<p>1. Определите этапы процесса оценки PLS</p> <table border="1"> <thead> <tr> <th>Буква</th> <th>Этап</th> </tr> </thead> <tbody> <tr> <td>А</td> <td>Оценка измерительной модели: расчёт loadings, CR, AVE, HTMT</td> </tr> <tr> <td>В</td> <td>Спецификация концептуальной рамки и формулировка гипотез</td> </tr> <tr> <td>С</td> <td>Интерпретация результатов и формирование бизнес-рекомендаций (IPMA, executive summary)</td> </tr> <tr> <td>Д</td> <td>Тестирование гипотез с помощью бутстрэппинга (<math>\geq 5000</math> итераций), расчёт медиации/модерации</td> </tr> <tr> <td>Е</td> <td>Оценка структурной модели: коэффициенты путей, <math>R^2</math>, <math>f^2</math>, <math>Q^2</math>, SRMR</td> </tr> <tr> <td>Ф</td> <td>Подготовка данных: обработка пропусков, выбросов,</td> </tr> </tbody> </table>	Буква	Этап	А	Оценка измерительной модели: расчёт loadings, CR, AVE, HTMT	В	Спецификация концептуальной рамки и формулировка гипотез	С	Интерпретация результатов и формирование бизнес-рекомендаций (IPMA, executive summary)	Д	Тестирование гипотез с помощью бутстрэппинга ( $\geq 5000$ итераций), расчёт медиации/модерации	Е	Оценка структурной модели: коэффициенты путей, $R^2$ , $f^2$ , $Q^2$ , SRMR	Ф	Подготовка данных: обработка пропусков, выбросов,
Буква	Этап															
А	Оценка измерительной модели: расчёт loadings, CR, AVE, HTMT															
В	Спецификация концептуальной рамки и формулировка гипотез															
С	Интерпретация результатов и формирование бизнес-рекомендаций (IPMA, executive summary)															
Д	Тестирование гипотез с помощью бутстрэппинга ( $\geq 5000$ итераций), расчёт медиации/модерации															
Е	Оценка структурной модели: коэффициенты путей, $R^2$ , $f^2$ , $Q^2$ , SRMR															
Ф	Подготовка данных: обработка пропусков, выбросов,															

		проверка мультиколлинеарности														
		2. Определите последовательность этапов EDA														
		<table border="1"> <thead> <tr> <th>Буква</th> <th>Этап</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>Оценка измерительной модели: расчёт loadings, CR, AVE, HTMT</td> </tr> <tr> <td>B</td> <td>Спецификация концептуальной рамки и формулировка гипотез</td> </tr> <tr> <td>C</td> <td>Интерпретация результатов и формирование бизнес-рекомендаций (IPMA, executive summary)</td> </tr> <tr> <td>D</td> <td>Тестирование гипотез с помощью бутстрэппинга (<math>\geq 5000</math> итераций), расчёт медиации/модерации</td> </tr> <tr> <td>E</td> <td>Оценка структурной модели: коэффициенты путей, <math>R^2</math>, <math>f^2</math>, <math>Q^2</math>, SRMR</td> </tr> <tr> <td>F</td> <td>Подготовка данных: обработка пропусков, выбросов, проверка мультиколлинеарности</td> </tr> </tbody> </table>	Буква	Этап	A	Оценка измерительной модели: расчёт loadings, CR, AVE, HTMT	B	Спецификация концептуальной рамки и формулировка гипотез	C	Интерпретация результатов и формирование бизнес-рекомендаций (IPMA, executive summary)	D	Тестирование гипотез с помощью бутстрэппинга ( $\geq 5000$ итераций), расчёт медиации/модерации	E	Оценка структурной модели: коэффициенты путей, $R^2$ , $f^2$ , $Q^2$ , SRMR	F	Подготовка данных: обработка пропусков, выбросов, проверка мультиколлинеарности
Буква	Этап															
A	Оценка измерительной модели: расчёт loadings, CR, AVE, HTMT															
B	Спецификация концептуальной рамки и формулировка гипотез															
C	Интерпретация результатов и формирование бизнес-рекомендаций (IPMA, executive summary)															
D	Тестирование гипотез с помощью бутстрэппинга ( $\geq 5000$ итераций), расчёт медиации/модерации															
E	Оценка структурной модели: коэффициенты путей, $R^2$ , $f^2$ , $Q^2$ , SRMR															
F	Подготовка данных: обработка пропусков, выбросов, проверка мультиколлинеарности															
Задание комбинированного типа с выбором одного правильного ответа из предложенных и обоснованием выбора	<p>1. Внимательно прочитать текст задания и понять, что в качестве ответа ожидается только один из предложенных вариантов.</p> <p>2. Внимательно прочитать предложенные варианты ответа.</p> <p>3. Выбрать один верный ответ.</p> <p>4. Записать только номер (или букву) выбранного варианта ответа.</p> <p>5. Записать аргументы, обосновывающие выбор ответа (например, 4 текст обоснования).</p>	<p>Какой механизм пропусков является НАИБОЛЕЕ проблематичным для анализа и требует специализированных методов?</p> <p>A) MCAR (Missing Completely At Random)</p> <p>B) MAR (Missing At Random)</p> <p>C) MNAR (Missing Not At Random) — вероятность пропуска зависит от самого ненаблюдаемого значения</p> <p>D) Все механизмы одинаково просты в обработке</p> <p>2. Какая метрика центральности НАИЛУЧШЕ выявляет узлы, контролирующие потоки информации между сообществами?</p> <p>A) Degree centrality (степень узла)</p> <p>B) Betweenness centrality (центральность по посредничеству)</p> <p>C) Closeness centrality (центральность по близости)</p> <p>D) Eigenvector centrality (центральность по собственному вектору)</p>														
Задание открытого типа с развернутым ответом	<p>1. Внимательно прочитать текст задания и понять суть вопроса.</p> <p>2. Продумать логику и полноту ответа.</p> <p>3. Записать ответ, используя четкие компактные формулировки.</p> <p>4. В случае расчетной задачи, записать решение и ответ</p>	<p>1. Какой тип пропусков (MCAR, MAR, MNAR) наиболее проблематичен для анализа и почему?</p> <p>2. Что показывает график qq-plot и как его использовать для проверки нормальности?</p>														

### 6.3. Критерии и шкала оценивания на основе БРС.

*Критерии и балльная шкала определены в БРС. В moodle оценка выставляется по результатам одной попытки. На попытку дается 60 минут. В тест включены 30 вопросов из разных мет учебной дисциплины.*

Максимальное число баллов на зачет – 40. Из них тридцать баллов за тест и 10 баллов за активность на занятиях слушателей или наличие сертификата по одному из курсов, посвященных предсказательной аналитике.

6.4. Описание дополнительных материалов и оборудования, необходимых для выполнения проверочных заданий (*при необходимости*).

Для решения задач открытого типа (кейсов, ПКЗ, ПИЗ), тестовых заданий студенту разрешается использование калькулятора; программ для работы с электронными таблицами для обработки, анализа и визуализации данных. Для построения интеллект-карты и моделей в различных нотациях студенту можно использовать любой соответствующий онлайн-инструмент.

## **7. Методические материалы по освоению дисциплины (модуля)**

Для закрепления изученного материала даны вопросы по каждой теме дисциплины, на которые следует самостоятельно найти ответы, основные материалы размещены в moodle.

Важной составной частью учебного процесса в вузе являются практические занятия. Практические занятия проводятся главным образом по дисциплинам, требующим закрепления навыков решения задач, и помогают студентам глубже усвоить учебный материал, приобрести умения применять принципы системного подхода к решению разнообразных задач, определять и оценивать ресурсы и существующие ограничения разного рода проектов.

Необходимо помнить, что на лекции обычно рассматривается не весь материал, а только его часть. Остальная его часть восполняется в процессе самостоятельной работы. В связи с этим работа с рекомендованной литературой обязательна.

После изучения базовых тем курса проводится текущий контроль знаний студентов в виде опроса или письменного тестирования. Типовые тесты и задания по темам дисциплины приведены в специальном разделе данной рабочей программы.

## **8. Учебная литература и ресурсы информационно-телекоммуникационной сети Интернет**

### **8.1. Основная литература**

1. Алламмор Джей, Гроотендорст Маартен. Большие языковые модели на практике. Понимание языка и генерация текстов. – Астана, Алист, 2026. -416 с.
2. Анализ данных : учебник для вузов / под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2026. — 448 с. — (Высшее образование). — ISBN 978-5-534-19964-2. — Текст : электронный // Образовательная

платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/583032> (дата обращения: 01.05.2026).

1. Афанасьев, Владимир Николаевич. Анализ временных рядов и прогнозирование. - Саратов: Ай Пи Ар Медиа. – 310 с. Текст : электронный. - URL: <http://www.iprbookshop.ru/90196.html> (дата обращения: 12.11.2020). - Режим доступа: для авторизир. Пользователей

3. Голоктионова Ю.Г., Ильминская С.А., Илюхина И.Б., Луговской А.М., Лисичкина Н.В. и др. Прогнозирование и планирование в экономике. - Москва: Прометей – 544 с. Текст : электронный. - URL: <http://www.iprbookshop.ru/94511.html> (дата обращения: 01.10.2020). - Режим доступа: для авторизир. пользователей

4. Миркин, Б. Г. Базовые методы анализа данных : учебник и практикум для вузов / Б. Г. Миркин. — 3-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2026. — 297 с. — (Высшее образование). — ISBN 978-5-534-19709-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/583143> (дата обращения: 01.05.2026).

5. Наумов, В. Н. Методы прогнозирования временных рядов : учебное пособие для вузов / В. Н. Наумов. — 2-е изд., стер. — Санкт-Петербург : Лань, 2025. — 196 с. — ISBN 978-5-507-53013-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/464216> (дата обращения: 01.05.2026). — Режим доступа: для авториз. пользователей.

6. О'Нил, Кэти. Data Science : Инсайдерская информация для новичков. Включая язык R : [пер. с англ.] – СПб. Питер. – 368 с. Текст: электронный. - URL: <http://new.ibooks.ru/bookshelf/359209/reading> (дата обращения: 25.01.2021)

7. Рон Хайндман и Джордж Атанасопулос X15 Прогнозирование: принципы и практика / пер. с англ. А. В. Логунова. – М.: ДМК Пресс, 2023. – 458 с.: ил.

## 8.2. Дополнительная литература

1. Principles of Econometrics with R [Электронный ресурс] – URL: <https://bookdown.org/ccolonescu/RPoE4/>

2. Балдин К.В. Эконометрика : учеб. пособие для студентов вузов, обучающихся по специальностям "Финансы и кредит", "Бухгалтерский учет, анализ и аудит" / К.В. Балдин, О.Ф. Быстров, М.М. Соколов. - М. : ЮНИТИ-ДАНА, 2004. - 254с.

3. Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Прогнозирование временных рядов и процессов. 3-е изд. - СПб. : БХВ-Петербург, 2010. - 512 с.

4. Боровиков В. П. Прогнозирование в системе Statistica в среде Windows : основы теории и интенсивная практика на компьютере : учеб. пособие для вузов, рек. М-вом образования Рос. Федерации / В. П. Боровиков, Г. И. Ивченко. - М. : Финансы и статистика, 2000. - 378 с.

5. Буре В. М. Основы эконометрики: учеб. пособие / В. М. Буре, Е. А. Евсеев.. — СПб.: Изд-во С.-Петерб. ун-та, 2004. - 72 с.
6. Жерон, Орельен. ПрМ.:Диалектика. -684 с. Прикладное машинное обучение с помощью Scikit-Learn и TensorFlow : Концепции, инструменты и техники для создания интеллектуальных систем : полноцветное издание : перевод с английского – М.: Диалектика, 2018 – 688 с.
7. Наумов В.Н. Анализ данных и машинное обучение: методы и инструментальные средства. Федер. гос. бюджет. образоват. учреждение высш. образования "Рос. акад. нар. хоз-ва и гос. службы при Президенте Рос. Федерации", Сев.-Зап. ин-т упр. - СПб. : СЗИУ - фил. РАНХиГС, 2020. - 260 с.
8. Наумов, Владимир Николаевич. Средства бизнес-аналитики: учеб. пособие / В. Н. Наумов ; Федер. гос. бюджет. образоват. учреждение высш. образования "Рос. акад. нар. хоз-ва и гос. службы при Президенте Рос. Федерации", Сев.-Зап. ин-т упр. - СПб. : СЗИУ - фил. РАНХиГС, 2016. - 107 с.
9. Нильсен Эйлин. Практический анализ временных рядов: прогнозирование со статистикой и машинное обучение. –М.: ООО Диалектика – 2021 – 544 с.
10. Саати Т. Л. Принятие решений при зависимостях и обратных связях : аналитические сети = Decision making with dependence and feedback : analytic network process / Т. Л. Саати ; пер. с англ. О. Н. Андрейчиковой ; науч. ред.: А. В. Андрейчиков, О. Н. Андрейчикова. - Изд. 2-е. - М. : УРСС, 2009. - 357 с.
11. Цыгичко В. Н. Прогнозирование социально-экономических процессов / В.Н. Цыгичко ; с предисл. Д.М. Гвишиани. - изд. 2-е, перераб. и доп. - М. : КомКнига, 2007. - 238 с.
12. Эконометрика : учебник / [И.И. Елисеева и др.] ; под ред. И.И. Елисевой. - М. : Проспект, 2009. - 288 с.
13. Эконометрика: учебник / В. С. Мхитарян [и др.] ; под ред. В. С. Мхитаряна. - М. : Проспект, 2009. - 380 с.

### **8.3 Нормативные правовые документы и иная правовая информация**

Не используются

#### **8.4 Интернет-ресурсы**

Обучающимся обеспечен доступ к материалам курса в СДО Академии <http://lms.ranepa.ru>, а также через сайт научной библиотеки к следующим подписным электронным ресурсам:

##### ***Русскоязычные ресурсы***

- Электронные учебники электронно-библиотечной системы (ЭБС) «Айбукс»
- Электронные учебники электронно-библиотечной системы (ЭБС) «Юрайт»
- Электронные учебники электронно-библиотечной системы (ЭБС) «Лань»

- Электронные учебники электронно-библиотечной системы (ЭБС) «ZNANI-UM.COM»
- Электронные учебники электронно-библиотечной системы (ЭБС) «BOOK.RU»
- Электронные учебники электронно-библиотечной системы (ЭБС) «IPR SMART»

### **9. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы**

№ п/п	Наименование
1.	Специализированные залы для проведения лекций, оснащенные персональным компьютером/ноутбуком и мультимедийным проектором
2.	Аудитории и компьютерные классы, оборудованные посадочными местами и персональными компьютерами с выходом в Интернет для проведения практических занятий
3.	«МТС Линк» — российская платформа для онлайн-коммуникаций и совместной работы команд ; «Яндекс Телемост» — сервис для видеоконференций от Яндекса; Я-мессенджер
4.	Технические средства обучения: персональные компьютеры; программные средства, обеспечивающие просмотр видеофайлов в форматах AVI, MPEG-4, DivX, RMVB, WMV; программы для работы с электронными таблицами для обработки, анализа и визуализации данных; соответствующие онлайн-инструменты для построения интеллект-карты и моделей в различных нотациях
5.	Научная библиотека (в т.ч. электронные информационные ресурсы научной библиотеки)
6.	СДО Академии <a href="https://lms.ranepa.ru/">https://lms.ranepa.ru/</a>
7.	R, Rstudio, Anaconda Navigator, VS code, JASP, jamovi, excel