

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Панарин Андрей Александрович

Должность: Ректор

Дата подписания: 18.06.2026 12:03:06

Уникальный программный ключ:

a5da3d9896e9d535580e3f9a7da4892154ef8502



Образовательная автономная некоммерческая организация высшего образования

«Московский психолого-социальный университет»

Лицензия № 1478 от 28 мая 2015 г., серия 90Л01 № 0008476 (бессрочная)

Свидетельство государственной аккредитации № 2783 от 07 марта 2018 года, серия 90А01 №0002920 (бессрочно)

Кафедра «Экономики и цифровых технологий»

Фонд

оценочных средств

по дисциплине: «Методы интеллектуального анализа данных»

Уровень высшего образования

БАКАЛАВРИАТ

Направление подготовки

38.03.05 Бизнес-информатика

Направленность (профиль)

Data Science

Квалификация (степень) выпускника

бакалавр

Форма обучения

очная

Москва

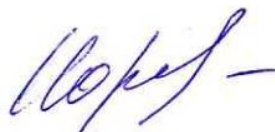
2026 год набора

Фонд оценочных средств предназначен для контроля знаний обучающихся по направлению подготовки 38.03.05 Бизнес-информатика, утвержденным приказом Министерство науки и высшего образования Российской Федерации от 29 июля 2020г. № 838» по дисциплине «Методы интеллектуального анализа данных».

Фонд оценочных средств рассмотрен и утвержден на заседании кафедры «Экономики и цифровых технологий»

Протокол № 7 от 22 января 2026 г.

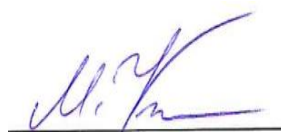
И.о. заведующий кафедрой
«Экономики и цифровых технологий»



И.Ф. Иорданиди

Согласовано:

Декан экономического факультета



М.К. Чистякова

1. Перечень планируемых результатов обучения, соотнесенных с планируемыми результатами освоения основной профессиональной образовательной программы УК-1.1, ПК-1.2; ПК-3.1.

Код и описание компетенции	Код и наименование индикатора достижения	Планируемые результаты обучения по дисциплине
УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1 Осуществляет поиск необходимой информации, опираясь на результаты анализа поставленной задачи.	<u>Контактная работа:</u> Лекции Практические занятия Самостоятельная работа
ПК-1. Способен осуществлять концептуально-логическое проектирование ИТ-сервиса, автоматизированной информационной системы, системы управления, программного, информационному продукту или средству и сопровождение разработанных проектных решений	ПК-1.2 Осуществляет концептуально-логическое проектирование к ИТ-сервису, автоматизированной информационной системе, системе управления, программного, информационному продукту или средству	<u>Контактная работа:</u> Лекции Практические занятия Самостоятельная работа
ПК – 3 Способен формировать возможные решения на основе разработанных для них целевых показателей	ПК-3.1. Собирает и использует информацию бизнес-анализа для формирования возможных решений.	<u>Контактная работа:</u> Лекции Практические занятия Самостоятельная работа

ТИПОВЫЕ ИНСТРУКЦИИ ПО ВЫПОЛНЕНИЮ ТЕСТОВЫХ ЗАДАНИЙ

Тип задания	Инструкция
Задание закрытого типа с выбором одного или нескольких ответов	Прочитайте текст и выберите правильный ответ (Если несколько ответов, то прочитайте текст и выберите правильные ответы)
Задание закрытого типа на установление соответствия	Прочитайте текст и установите соответствие
Задания закрытого типа на установление правильной последовательности	Прочитайте текст и установите последовательность
Задания комбинированного типа с выбором одного верного ответа с обоснованием	Прочитайте текст, выберите правильный ответ и запишите аргументы, обосновывающие выбор ответа
Задания комбинированного типа с выбором нескольких ответов с обоснованием	Прочитайте текст, выберите правильные ответы и запишите аргументы, обосновывающие выбор ответов
Задания с развернутым ответом	Прочитайте текст и запишите развернутый

	обоснованный ответ
--	---------------------------

СИСТЕМА ОЦЕНИВАНИЯ ЗАДАНИЙ

Тип задания	Указания по оцениванию	Результат оценивания
Задание закрытого типа на установление соответствия	Задание закрытого типа на установление соответствия считается верным если правильно установлены все соответствия	Верно/неверно
Задания закрытого типа на установление правильной последовательности	Задание закрытого типа на установление правильной последовательности считается верным если правильно указываются все последовательности	Верно/неверно
Задания комбинированного типа с выбором одного верного ответа с обоснованием	Задание комбинированного типа с выбором одного верного ответа из предложенных с обоснованием считается верным если правильно указан ответ и приведены корректные аргументы, используемые при выборе ответа.	Верно/неверно
Задания комбинированного типа с выбором нескольких ответов с обоснованием	Задание комбинированного типа с выбором нескольких ответов из предложенных с обоснованием считается верным если правильно указаны ответы и приведены корректные аргументы, используемые при выборе ответа.	Верно/неверно
Задания открытого типа с развернутым ответом	Задания открытого типа с развернутым ответом считается верным, если ответ совпадает с эталоном по содержанию и полноте.	Верно/неверно

3. Контрольные задания или иные материалы, необходимые для процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения основной профессиональной образовательной программы

3.1. Задания для проведения текущего контроля обучающихся

Содержание вопроса	Компетенции	Уровень освоения
<i>Прочитайте текст и выберите правильный ответ</i> 1. Что такое интеллектуальный анализ данных (Data Mining)? а) Ручной анализ больших объемов данных.	УК-1.1 Осуществляет поиск необходимой информации, опираясь на результаты анализа	Базовый 1-3 минуты

b) Извлечение значимой информации из больших объемов данных с помощью автоматизированных методов. c) Хранение данных в базе данных. d) Визуализация данных. 2. Основная цель интеллектуального анализа данных: a) Создание красивых графиков. b) Извлечение скрытых закономерностей и предсказание будущих событий. c) Хранение данных. d) Обработка данных. 3. Какие типы данных используются в Data Mining? a) Только структурированные. b) Только неструктурированные. c) Структурированные, неструктурированные, полуструктурированные. d) Только числовые. 4. Что такое структурированные данные? a) Не имеет значения. b) Данные, организованные в таблицы с строками и столбцами. c) Текст, изображения, аудио. d) Данные в неформализованном виде. 5. Что такое неструктурированные данные? a) Не имеет значения. b) Текст, изображения, аудио, видео. c) Данные в таблицах. d) Данные в формализованном виде. 6. Что такое полуструктурированные данные? a) Не имеет значения. b) Данные, имеющие частичную структуру. Например, XML файлы. c) Данные в таблицах. d) Данные в формализованном виде. 7. Какие этапы интеллектуального анализа данных существуют? a) Отсутствуют этапы. b) Сбор данных, предварительная обработка, выбор метода, моделирование, интерпретация результатов. c) Только сбор данных. d) Только интерпретация результатов.	поставленной задачи.	
Прочитайте текст и выберите правильный ответ 8. Что такое предварительная обработка данных? a) Не имеет значения. b) Очистка, преобразование и подготовка данных к анализу.	ПК-1.2 Осуществляет концептуально-логическое проектирование к ИТ-	Повышенный 3-5 минут

c) Только для структурированных данных. d) Только для неструктурированных данных. 9. Какие методы предварительной обработки данных существуют? a) Отсутствуют методы. b) Обработка пропущенных значений, нормализация, дискретизация. c) Только обработка пропущенных значений. d) Только нормализация. 10. Что такое кластеризация? a) Не имеет значения. b) Разбиение данных на группы (кластеры) с похожими характеристиками. c) Только для структурированных данных. d) Только для неструктурированных данных. 11. Что такое k-means? a) Не имеет значения. b) Алгоритм кластеризации. c) Только для структурированных данных. d) Только для неструктурированных данных. 12. Что такое иерархическая кластеризация? a) Не имеет значения. b) Алгоритм кластеризации, строящий иерархическое дерево кластеров. c) Только для структурированных данных. d) Только для неструктурированных данных. 13. Что такое классификация? a) Не имеет значения. b) Приписывание данным к заранее известным классам. c) Только для структурированных данных. d) Только для неструктурированных данных.	сервису, автоматизированной информационной системе, системе управления, программному, информационному продукту или средству	
Прочитайте текст и выберите правильный ответ 14. Что такое регрессия? a) Не имеет значения. b) Предсказание непрерывной величины. c) Только для структурированных данных. d) Только для неструктурированных данных. 15. Что такое линейная регрессия? a) Не имеет значения. b) Метод регрессии, использующий линейную модель.	ПК-3.1. Собирает и использует информацию бизнес-анализа для формирования возможных решений	

c) Только для структурированных данных. d) Только для неструктурированных данных. 16. Что такое логистическая регрессия? a) Не имеет значения. b) Метод классификации, использующий логистическую функцию. c) Только для структурированных данных. d) Только для неструктурированных данных. 17. Что такое дерево решений? a) Не имеет значения. b) Метод классификации и регрессии, представляющий решение в виде дерева. c) Только для структурированных данных. d) Только для неструктурированных данных. 18. Что такое нейронные сети? a) Не имеет значения. b) Модели, имитирующие работу человеческого мозга. c) Только для структурированных данных. d) Только для неструктурированных данных. 19. Что такое машинное обучение? a) Не имеет значения. b) Разработка алгоритмов, позволяющих компьютерам учиться на данных. c) Только для структурированных данных. d) Только для неструктурированных данных. 20. Что такое глубокое обучение? a) Не имеет значения. b) Подмножество машинного обучения, использующее многослойные нейронные сети. c) Только для структурированных данных. d) Только для неструктурированных данных.		
Прочитайте текст и запишите развернутый обоснованный ответ Задача 1: Кластеризация клиентов Условие: Компания имеет базу данных о своих клиентах, включающую информацию о возрасте, доходе и частоте покупок. Необходимо разделить клиентов на группы (кластеры) с похожими характеристиками. Какой метод	УК-1.1 Осуществляет поиск необходимой информации, опираясь на результаты анализа поставленной задачи. ПК-1.2 Осуществляет концептуально-	Высокий 5-10 минут

кластеризации лучше всего подходит для этой задачи? Обоснуйте свой выбор. Ответ: Для этой задачи хорошо подходит метод k-means. Он относительно прост в реализации и эффективен для больших наборов данных. Данные о возрасте, доходе и частоте покупок являются числовыми, что делает k-means подходящим выбором. Иерархическая кластеризация также может быть применена, но k-means обычно быстрее и эффективнее для больших объемов данных. Задача 2: Предсказание оттока клиентов Условие: Телекоммуникационная компания хочет предсказать, какие клиенты с большой вероятностью откажутся от услуг. Какие методы интеллектуального анализа данных можно использовать для решения этой задачи? Ответ: Для предсказания оттока клиентов можно использовать методы классификации. Подходящие методы включают логистическую регрессию, дерево решений, нейронные сети. Для начала необходимо подготовить обучающий набор данных, где каждый клиент помечен как "отток" или "нет оттока", а также содержащий релевантные характеристики клиентов (например, длительность пользования услугами, количество звонков в службу поддержки, средний счет). Задача 3: Рекомендательная система Условие: Онлайн-магазин хочет создать рекомендательную систему, предлагающую клиентам товары, которые им могут понравиться. Какие методы можно использовать? Ответ: Для создания рекомендательной системы можно использовать методы коллаборативной фильтрации (рекомендации на основе предпочтений других пользователей с похожими вкусами) и/или контент-based фильтрации (рекомендации на основе характеристик товара). Алгоритмы коллаборативной фильтрации включают user-based и item-based подходы. Контент-based фильтрация использует информацию о товарах (жанр, автор, описание). Задача 4: Анализ текстовых данных Условие: Компания хочет проанализировать отзывы клиентов о своем продукте, извлекая из них общие темы и настроения. Какие методы можно использовать?	логическое проектирование к ИТ-сервису, автоматизированной информационной системе, системе управления, программному, информационному продукту или средству ПК-3.1. Собирает и использует информацию бизнес-анализа для формирования возможных решений	
--	---	--

Ответ: Для анализа текстовых данных можно использовать методы обработки естественного языка (NLP). Это включает токенизацию (разбиение текста на слова), лемматизацию/стемминг (приведение слов к их основным формам), извлечение сущностей, анализ настроений (определение положительных, отрицательных или нейтральных отзывов). Для темы используются методы топик моделирования (например, LDA).		
---	--	--

3.2.2. Задания для проведения промежуточной аттестации обучающихся

Содержание вопроса	Компетенции	Уровень освоения
<i>Прочитайте текст и выберите правильный ответ</i> 21. Что такое ассоциативные правила? а) Не имеет значения. б) Правила, описывающие связи между переменными в данных. в) Только для структурированных данных. г) Только для неструктурированных данных. 22. Что такое алгоритм Apriori? а) Не имеет значения. б) Алгоритм для поиска ассоциативных правил. в) Только для структурированных данных. г) Только для неструктурированных данных. 23. Что такое подготовка данных? а) Не имеет значения. б) Очистка, преобразование и подготовка данных к анализу. в) Только для структурированных данных. г) Только для неструктурированных данных. 24. Что такое выбросы? а) Не имеет значение. б) Значения, значительно отличающиеся от других данных. в) Только для числовых данных. г) Только для категориальных данных. 25. Что такое нормализация данных? а) Не имеет значение. б) Приведение данных к единому масштабу. в) Только для числовых данных.	УК-1.1 Осуществляет поиск необходимой информации, опираясь на результаты анализа поставленной задачи.	Базовый 1-3 минуты

d) Только для категориальных данных. 26. Что такое дискретизация данных? a) Не имеет значение. b) Разбиение непрерывных данных на интервалы. c) Только для числовых данных. d) Только для категориальных данных. 27. Что такое оценка модели? a) Не имеет значения. b) Оценка точности и качества модели. c) Только для классификации. d) Только для регрессии.		
Прочитайте текст и выберите правильный ответ 28. Что такое точность (precision)? a) Не имеет значения. b) Доля правильно классифицированных объектов от всех объектов, отнесенных к данному классу. c) Только для классификации. d) Только для регрессии. 29. Что такое полнота (recall)? a) Не имеет значения. b) Доля правильно классифицированных объектов от всех объектов данного класса. c) Только для классификации. d) Только для регрессии. 30. Что такое F1-мера? a) Не имеет значения. b) Гармоническое среднее точности и полноты. c) Только для классификации. d) Только для регрессии. 31. Что такое RMSE (Root Mean Squared Error)? a) Не имеет значения. b) Квадратный корень из средней квадратичной ошибки. c) Только для регрессии. d) Только для классификации. 32. Что такое R-квадрат? a) Не имеет значения. b) Коэффициент детерминации. c) Только для регрессии. d) Только для классификации.	ПК-1.2 Осуществляет концептуально-логическое проектирование к ИТ-сервису, автоматизированной информационной системе, системе управления, программному, информационному продукту или средству	Повышенный 3-5 минут

33. Что такое кросс-валидация? a) Не имеет значения. b) Метод оценки модели с помощью разбиения данных на несколько подвыборок. c) Только для классификации. d) Только для регрессии.		
Прочитайте текст и выберите правильный ответ 34. Что такое тестирование гипотез? a) Не имеет значения. b) Проверка статистических гипотез о данных. c) Только для структурированных данных. d) Только для неструктурированных данных. 35. Что такое выборка? a) Не имеет значения. b) Подмножество данных, используемое для анализа. c) Только для структурированных данных. d) Только для неструктурированных данных. 36. Что такое выборочное среднее? a) Не имеет значения. b) Среднее арифметическое значений в выборке. c) Только для числовых данных. d) Только для категориальных данных. 37. Что такое выборочная дисперсия? a) Не имеет значения. b) Мера рассеивания значений в выборке. c) Только для числовых данных. d) Только для категориальных данных. 38. Что такое доверительный интервал? a) Не имеет значения. b) Интервал, в котором с определенной вероятностью находится истинное значение параметра. c) Только для числовых данных. d) Только для категориальных данных. 39. Что такое р-значение? a) Не имеет значения. b) Вероятность получить результаты не менее экстремальные, чем наблюдаемые, если нулевая гипотеза верна. c) Только для тестирования гипотез. d) Только для оценки моделей.	ПК-3.1. Собирает и использует информацию бизнес-анализа для формирования возможных решений	

40. Что такое визуализация данных? a) Не имеет значения. b) Представление данных в графическом виде. c) Только для структурированных данных. d) Только для неструктурированных данных.		
Прочитайте текст и запишите развернутый обоснованный ответ Задача 5: Выявление аномалий Условие: Компания заметила необычное поведение в своей системе платежей. Как обнаружить аномалии? Ответ: Для обнаружения аномалий можно использовать методы анализа выбросов. Это включает статистические методы (например, определение выбросов по стандартному отклонению), методы на основе плотности (например, LOF – Local Outlier Factor), а также нейронные сети и машинное обучение. Задача 6: Предсказание продаж Условие: Магазин хочет предсказать свои продажи на следующий месяц. Какие методы подходят? Ответ: Для предсказания продаж можно использовать методы регрессии, такие как линейная регрессия, регрессия с временными рядами (ARIMA, Prophet), нейронные сети (RNN). Необходимо подготовить исторические данные о продажах, возможно, учитывая внешние факторы (например, праздники, сезонность). Задача 7: Оценка качества модели Условие: Разработана модель классификации, предсказывающая удовлетворенность клиентов. Как оценить ее качество? Ответ: Для оценки качества модели классификации можно использовать метрики, такие как точность (precision), полнота (recall), F1-мера, AUC-ROC (Area Under the ROC Curve). Также необходимо использовать кросс-валидацию для получения более надежной оценки обобщающей способности модели. Задача 8: Выбор метода для задачи	УК-1.1 Осуществляет поиск необходимой информации, опираясь на результаты анализа поставленной задачи. ПК-1.2 Осуществляет концептуально-логическое проектирование к ИТ-сервису, автоматизированной информационной системе, системе управления, программному, информационному продукту или средству ПК-3.1. Собирает и использует информацию бизнес-анализа для формирования возможных решений	Высокий 5-10 минут

Условие: Перед вами стоит задача сегментации клиентов по их поведению на сайте (просмотренные страницы, время, проведенное на сайте, покупки). Какой метод интеллектуального анализа данных вы бы выбрали? Ответ: Для сегментации клиентов по их поведению на сайте наилучшим образом подходит метод кластеризации. В данном случае можно использовать k-means или иерархическую кластеризацию в зависимости от размера набора данных и желаемой структуры кластеров. Предварительная обработка данных (нормализация, обработка пропущенных значений) является важным этапом		
--	--	--

3.3. Вопросы к зачету (промежуточная аттестация), формирование компетенций УК-1.1, ПК-1.2; ПК-3.1.

1. Эволюция методов обработки и анализа данных.
2. Переход от статистической обработки к методам OLAP и Data Mining. Принципы, лежащие в основе методов OLAP и Data Mining.
3. Общая характеристика методов интеллектуального анализа данных. Обзор основных задач, решаемых методами Data Mining.
4. Алгоритмы предварительной обработки данных. Очистка данных: факторный анализ, корреляционный анализ, фильтрация.
5. Трансформация данных: метод скользящего окна, квантование, сортировка и группировка, слияние данных.
6. Понятие и основные принципы «извлечения знаний». Основные задачи, решаемые методами Data Mining.
7. Автокорреляция. Линейная и логистическая регрессия.
8. Дерево решений: принципы построения, использование, интерпретация результатов.
9. Метод поиска ассоциативных правил, области его применения.
10. Нейросетевые методы анализа данных. Принципы функционирования искусственных нейронных сетей. Основные области их применения.
11. Многослойные нейросети. Аппроксимация функциональных зависимостей. Прогнозирование временных рядов.
12. Карты Кохонена. Задачи кластеризации и визуализации больших объемов данных. Интерпретация результатов.
13. Инструменты многомерного статистического анализа. Статистический анализ данных. Основные возможности, спектр решаемых задач.
14. Табличные процессоры как инструмент анализа. Эволюция инструментов: от Excel до SPSS и Statistica.
15. Этапы выполнения многомерного статистического анализа и интерпретация результатов
16. Платформа Deductor Studio: основные возможности, сфера применения.
17. Этапы решения задач с помощью Deductor Studio. Импорт/экспорт данных. Визуализация результатов. Возможности автоматизации отдельных этапов анализа данных.
18. Экспертные системы. Составные части экспертной системы. Приобретение знаний. Извлечение знаний из данных. Участники процесса проектирования: эксперты, инженеры по знаниям, конечные пользователи.
19. Специфика анализа данных в менеджменте. Отбор данных для анализа.

Предварительная подготовка данных. Понижение размерности, сглаживание аномалий, фильтрация, группировка. Этапы проведения интеллектуального анализа.

Интерпретация результатов анализа

20. Принципы выбора методов и инструментов интеллектуального анализа .

21. Основные понятия машинного обучения. Основные постановки задач.

Примеры прикладных задач.

22. Линейные пространства. Векторы и матрицы. Линейная независимость.

Обратная матрица.

23. Производная и градиент функции. Градиентный спуск. Выпуклые функции.

24. Случайные величины. Дискретные и непрерывные распределения. Примеры.

25. Оценивание параметров распределений, метод максимального правдоподобия.

26. Бутстрэппинг.

27. Линейные методы классификации и регрессии: функционалы качества, методы настройки, особенности применения.

28. Метрики качества алгоритм регрессии и классификации.

29. Оценивание качества алгоритмов. Отложенная выборка, ее недостатки. Оценка полного скользящего контроля. Кросс-валидация. Leave-one-out.

30. Деревья решений. Методы построения деревьев. Их регуляризация.

31. Композиции алгоритмов. Разложение ошибки на смещение и разброс.

32. Случайный лес, его особенности.

33. Методы поиска выбросов в данных. Методы восстановления пропусков в данных.

Работа с несбалансированными выборками.

34. Задача анализа потребительской корзины. Поддержка и достоверность.

Частые, замкнутые и максимальные частые множества. Алгоритм Априори.

35. Задача кластеризации. Алгоритм K-Means. Оценки качества кластеризации

4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения основной профессиональной образовательной программы

№ п/п	Форма контроля/ коды оцениваемых компетенций	Процедура оценивания	Шкала и критерии оценки, балл
1.	Зачет УК-1.1, ПК-1.2., ПК-3.1.	При подготовке к зачету необходимо ориентироваться на конспекты лекций, рабочую программу дисциплины, нормативную, основную и дополнительную учебную литературу. Основное в подготовке к сдаче зачета - это повторение всего материала дисциплины. При подготовке к сдаче зачета обучающийся весь объем работы должен распределять равномерно по дням, отведенным для подготовки к зачету, контролировать каждый день выполнение намеченной работы. В период подготовки к зачету	1)«зачтено» - правильность ответов на вопросы билета (верное, четкое, достаточно глубокое изложение идей, понятий, фактов, нормативно-правового материала и т.п.) и правильное разрешение задачи; полнота и лаконичность ответа; степень использования и понимания научных и нормативных источников; умение связывать теорию с практикой; логика и аргументированность изложения материала; грамотное комментирование, приведение примеров, аналогий; культура речи;

		обучающийся вновь обращается к уже изученному (пройденному) учебному материалу. Подготовка обучающегося к зачету включает в себя три этапа: самостоятельная работа в течение семестра; непосредственная подготовка в дни, предшествующие зачету по темам курса; подготовка к ответу на задания, содержащиеся в вопросах (тестах) зачета. Зачет проводится по вопросам (тестам), охватывающим весь пройденный материал дисциплины, включая вопросы, отведенные для самостоятельного изучения.	2) «не зачтено» предполагает, что обучающимся либо не дан ответ на вопрос и (или) не решена предложенная задача, либо обучающийся не знает основных понятий, не может определить предмет дисциплины.
2.	Тестирование УК-1.1, ПК-1.2., ПК-3.1.	Полнота знаний теоретического контролируемого материала. Количество правильных ответов	«отлично» - процент правильных ответов = > 90%; «хорошо» - процент правильных ответов = > 70%; «удовлетворительно» - процент правильных ответов = > 50%; «неудовлетворительно» - процент правильных ответов < 50%.

С целью определения уровня овладения компетенциями, закрепленными за дисциплиной, в заданные преподавателем сроки проводится текущий и промежуточный контроль знаний, умений и навыков каждого обучающегося. Все виды текущего контроля осуществляются на практических занятиях. Исключение составляет устный опрос, который может проводиться в начале или конце лекции в течение 15-20 мин. с целью закрепления знаний терминологии по дисциплине. При оценке компетенций принимается во внимание формирование профессионального мировоззрения, определенного уровня включённости в занятия, рефлексивные навыки, владение изучаемым материалом.

Процедура оценивания компетенций обучающихся основана на следующих стандартах:

1. Периодичность проведения оценки.
2. Многоступенчатость: оценка (как преподавателем, так и обучающимися группы) и самооценка обучающегося, обсуждение результатов и комплекс мер по устранению недостатков.
3. Единство используемой технологии для всех обучающихся, выполнение условий сопоставимости результатов оценивания.
4. Соблюдение последовательности проведения оценки.

Текущая аттестация обучающихся. Текущая аттестация обучающихся по дисциплине проводится в соответствии с локальными нормативными актами ОАНО ВО МПСУ и является обязательной.

Текущая аттестация проводится в форме опроса и контрольных мероприятий по оцениванию фактических результатов обучения обучающихся и осуществляется преподавателем дисциплины.

Объектами оценивания выступают:

- 1) учебная дисциплина (активность на занятиях, своевременность выполнения различных видов заданий, посещаемость всех видов занятий по аттестуемой дисциплине);
- 2) степень усвоения теоретических знаний в качестве «ключей анализа»;

- 3) уровень овладения практическими умениями и навыками по всем видам учебной работы;
- 4) результаты самостоятельной работы (изучение книг из списка основной и дополнительной литературы).

Активность обучающегося на занятиях оценивается на основе выполненных обучающимся работ и заданий, предусмотренных данной рабочей программой дисциплины.

Кроме того, оценивание обучающегося проводится на текущем контроле по дисциплине. Оценивание обучающегося на контрольной неделе проводится преподавателем независимо от наличия или отсутствия обучающегося (по уважительной или неуважительной причине) на занятии. Оценка носит комплексный характер и учитывает достижения обучающегося по основным компонентам учебного процесса за текущий период.

Оценивание обучающегося носит комплексный характер и учитывает достижения обучающегося по основным компонентам учебного процесса за текущий период с выставлением оценок в ведомости.

Промежуточная аттестация обучающихся. Промежуточная аттестация проводится в соответствии с локальными нормативными актами ОАНО ВО «МПСУ» и является обязательной.

Промежуточная аттестация проводится в соответствии с учебным планом в виде **зачета** в период зачётно-экзаменационной сессии в соответствии с графиком проведения.

Обучающиеся допускаются к зачету в случае выполнения ими учебного плана по дисциплине: выполнения всех заданий и мероприятий, предусмотренных программой дисциплины.

Оценка знаний обучающегося на зачёте определяется его учебными достижениями и результатами текущего контроля знаний и выполнением им заданий.

Основой для определения оценки служит уровень усвоения обучающимися материала, предусмотренного данной рабочей программой дисциплины.

5. Методические указания для обучающихся по освоению дисциплины

Вид деятельности	Методические указания по организации деятельности студента
Лекция	Написание конспекта лекций: кратко, схематично, последовательно фиксировать основные положения, выводы, формулировки, обобщения; пометить важные мысли, выделять ключевые слова, термины. Проверка терминов, понятий с помощью энциклопедий, словарей, справочников с выписыванием толкований в тетрадь. Обозначить вопросы, термины, материал, который вызывает трудности, пометить и попытаться найти ответ в рекомендуемой литературе. Если самостоятельно не удастся разобраться в материале, необходимо сформулировать вопрос и задать преподавателю на консультации, на практическом занятии.
Практические занятия	Проработка рабочей программы, уделяя особое внимание целям и задачам, структуре и содержанию дисциплины. Конспектирование источников. Работа с конспектом лекций, подготовка ответов к контрольным вопросам, просмотр рекомендуемой литературы, работа с текстом. Прослушивание аудио- и видеозаписей по заданной теме, решение расчетно-графических заданий, решение задач по алгоритму и др.
Индивидуальные задания	Знакомство с основной и дополнительной литературой, включая справочные издания, зарубежные источники, конспект основных положений, терминов, сведений, требующихся для запоминания и являющихся основополагающими в этой теме. Составление аннотаций к прочитанным литературным источникам и др.
Самостоятельная работа	Самостоятельная работа проводится с целью: систематизации и закрепления полученных теоретических знаний и практических умений обучающихся; углубления и расширения теоретических знаний студентов; формирования умений использовать нормативную, правовую, справочную документацию, учебную и специальную литературу; развития познавательных способностей и

	активности обучающихся: творческой инициативы, самостоятельности, ответственности, организованности; формирование самостоятельности мышления, способностей к саморазвитию, совершенствованию и самоорганизации; формирования профессиональных компетенций; развитию исследовательских умений обучающихся. Формы и виды самостоятельной работы: чтение основной и дополнительной литературы – самостоятельное изучение материала по рекомендуемым литературным источникам; работа с библиотечным каталогом, самостоятельный подбор необходимой литературы; работа со словарем, справочником; поиск необходимой информации в сети Интернет; конспектирование источников; реферирование источников; составление аннотаций к прочитанным литературным источникам; составление рецензий и отзывов на прочитанный материал; составление обзора публикаций по теме; составление и разработка терминологического словаря; составление хронологической таблицы; составление библиографии (библиографической картотеки); подготовка к различным формам текущей и промежуточной аттестации (к тестированию, зачету, экзамену); выполнение домашних контрольных работ; самостоятельное выполнение практических заданий репродуктивного типа (ответы на вопросы, тесты; выполнение творческих заданий). Технология организации самостоятельной работы обучающихся включает использование информационных и материально-технических ресурсов образовательного учреждения: библиотеку с читальным залом, укомплектованную в соответствии с существующими нормами; учебно-методическую базу учебных кабинетов, лабораторий и зала кодификации; компьютерные классы с возможностью работы в сети Интернет; аудитории (классы) для консультационной деятельности; учебную и учебно-методическую литературу, разработанную с учетом увеличения доли самостоятельной работы студентов, и иные методические материалы. Перед выполнением обучающимися внеаудиторной самостоятельной работы преподаватель проводит консультирование по выполнению задания, который включает цель задания, его содержания, сроки выполнения, ориентировочный объем работы, основные требования к результатам работы, критерии оценки. Во время выполнения обучающимися внеаудиторной самостоятельной работы и при необходимости преподаватель может проводить индивидуальные и групповые консультации. Самостоятельная работа может осуществляться индивидуально или группами обучающихся в зависимости от цели, объема, конкретной тематики самостоятельной работы, уровня сложности, уровня умений обучающихся. Контроль самостоятельной работы предусматривает: • соотнесение содержания контроля с целями обучения; объективность контроля; • валидность контроля (соответствие предъявляемых заданий тому, что предполагается проверить); • дифференциацию контрольно-измерительных материалов. Формы контроля самостоятельной работы: • просмотр и проверка выполнения самостоятельной работы преподавателем; • организация самопроверки, • взаимопроверки выполненного задания в группе; обсуждение результатов выполненной работы на занятии; • проведение письменного опроса; • проведение устного опроса; • организация и проведение индивидуального собеседования; организация и проведение собеседования с группой;
--	--

	• защита отчетов о проделанной работе.
Опрос	Опрос - это средство контроля, организованное как специальная беседа преподавателя с обучающимся на темы, связанные с изучаемой дисциплиной, и рассчитанное на выявление объема знаний по определенному разделу, теме, проблеме и т.п. Проблематика, выносимая на опрос определена в заданиях для самостоятельной работы обучающегося, а также может определяться преподавателем, ведущим семинарские занятия. Во время проведения опроса обучающийся должен уметь обсудить с преподавателем соответствующую проблематику на уровне диалога.
Коллоквиум	Коллоквиум (от латинского colloquium – разговор, беседа) – одна из форм учебных занятий, беседа преподавателя с учащимися на определенную тему из учебной программы. Цель проведения коллоквиума состоит в выяснении уровня знаний, полученных учащимися в результате прослушивания лекций, посещения семинаров, а также в результате самостоятельного изучения материала. В рамках поставленной цели решаются следующие задачи: • выяснение качества и степени понимания учащимися лекционного материала; • развитие и закрепление навыков выражения учащимися своих мыслей; • расширение вариантов самостоятельной целенаправленной подготовки учащихся; • развитие навыков обобщения различных литературных источников; • предоставление возможности учащимся сопоставлять разные точки зрения по рассматриваемому вопросу. В результате проведения коллоквиума преподаватель должен иметь представление: • качества лекционного материала; • сильных и слабых сторонах своей методики чтения лекций; • сильных и слабых сторонах своей методики проведения семинарских занятий; • об уровне самостоятельной работы учащихся; • об умении обучающихся вести дискуссию и доказывать свою точку зрения; • степени эрудированности учащихся; • степени индивидуального освоения материала конкретными обучающимися. В результате проведения коллоквиума обучающийся должен иметь представление: • об уровне своих знаний по рассматриваемым вопросам в соответствии с требованиями преподавателя и относительно других студентов группы; • недостатках самостоятельной проработки материала; • своем умении излагать материал; • своем умении вести дискуссию и доказывать свою точку зрения. В зависимости от степени подготовки группы можно использовать разные подходы к проведению коллоквиума. В случае, если большинство группы с трудом воспринимает содержание лекций и на практических занятиях демонстрирует недостаточную способность активно оперировать со смысловыми единицами и терминологией курса, то коллоквиум можно разделить на две части. Сначала преподаватель излагает базовые понятия, содержащиеся в программе. Это должно занять не более четверти занятия. Остальные три четверти необходимо посвятить дискуссии, в ходе которой обучающиеся должны убедиться и, главное, убедить друг друга в обоснованности и доказательности полученного видения вопроса и его соответствия реальной практике. Если же

	преподаватель имеет дело с более подготовленной, самостоятельно думающей и активно усваивающей смысловые единицы и терминологию курса аудиторией, то коллоквиум необходимо провести так, чтобы сами обучающиеся сформулировали изложенные в программе понятия, высказали несовпадающие точки зрения и привели практические примеры. За преподавателем остается роль модератора (ведущего дискуссии), который в конце «лишь» суммирует совместно полученные результаты.
Тестирование	Контроль в виде тестов может использоваться после изучения каждой темы курса. Итоговое тестирование можно проводить в форме: • компьютерного тестирования, т.е. компьютер произвольно выбирает вопросы из базы данных по степени сложности; • письменных ответов, т.е. преподаватель задает вопрос и дает несколько вариантов ответа, а обучающийся на отдельном листе записывает номера вопросов и номера соответствующих ответов. Для достижения большей достоверности результатов тестирования следует строить текст так, чтобы у обучающихся было не более 40 – 50 секунд для ответа на один вопрос. Итоговый тест должен включать не менее 60 вопросов по всему курсу. Значит, итоговое тестирование займет целое занятие.
Подготовка к зачёту	При подготовке к зачёту необходимо ориентироваться на конспекты лекций, рекомендуемую литературу и др. Основное в подготовке к сдаче зачёта – это повторение всего материала дисциплины, по которому необходимо сдавать зачёт. При подготовке обучающийся весь объем работы должен распределять равномерно по дням, отведенным для подготовки, контролировать каждый день выполнение намеченной работы. Подготовка включает в себя три этапа: • самостоятельная работа в течение семестра; • непосредственная подготовка в дни, предшествующие зачёту по темам курса; • подготовка к ответу на задания, содержащиеся в билетах (тестах) зачёта. Для успешной сдачи зачёта обучающиеся должны принимать во внимание, что: • все основные вопросы, указанные в рабочей программе, нужно знать, понимать их смысл и уметь его разъяснить; • указанные в рабочей программе формируемые профессиональные компетенции в результате освоения дисциплины должны быть продемонстрированы студентом; • семинарские занятия способствуют получению более высокого уровня знаний и, как следствие, успешной сдаче зачёта; • готовиться к зачёту необходимо начинать с первой лекции и первого семинара.



Образовательная автономная некоммерческая организация высшего образования
«Московский психолого-социальный университет»

Лицензия № 1478 от 28 мая 2015 г., серия 90Л01 № 0008476 (бессрочная)

Свидетельство государственной аккредитации № 2783 от 07 марта 2018 года, серия 90А01 №0002920 (бессрочно)

КЛЮЧИ ПРАВИЛЬНЫХ ОТВЕТОВ К ФОНДУ ОЦЕНОЧНЫХ СРЕДСТВ

по дисциплине: «Методы интеллектуального анализа данных»

Уровень высшего образования
БАКАЛАВРИАТ

Направление подготовки
38.03.05 Бизнес-информатика

Направленность (профиль)
Data Science

Квалификация (степень) выпускника
бакалавр

Форма обучения
очная

Москва
2025 год набора

ОТВЕТЫ К ТЕСТУ:

1. b
2. b
3. c
4. b
5. b
6. b
7. b
8. b
9. b
10. b
11. b
12. b
13. b
14. b
15. b
16. b
17. b
18. b
19. b
20. b
21. b
22. b
23. b
24. b
25. b
26. b
27. b
28. b
29. b
30. b
31. b
32. b
33. b
34. b
35. b
36. b
37. b
38. b
39. b
40. b

КЕЙСЫ

Задача 1: Кластеризация клиентов

Условие: Компания имеет базу данных о своих клиентах, включающую информацию о возрасте, доходе и частоте покупок. Необходимо разделить клиентов на группы (кластеры) с похожими характеристиками. Какой метод кластеризации лучше всего подходит для этой задачи? Обоснуйте свой выбор.

Ответ: Для этой задачи хорошо подходит метод k-means. Он относительно прост в реализации и эффективен для больших наборов данных. Данные о возрасте, доходе и частоте покупок являются числовыми, что делает k-means подходящим выбором. Иерархическая кластеризация также может быть применена, но k-means обычно быстрее и эффективнее для больших объемов данных.

Задача 2: Предсказание оттока клиентов

Условие: Телекоммуникационная компания хочет предсказать, какие клиенты с большой вероятностью откажутся от услуг. Какие методы интеллектуального анализа данных можно использовать для решения этой задачи?

Ответ: Для предсказания оттока клиентов можно использовать методы классификации. Подходящие методы включают логистическую регрессию, дерево решений, нейронные сети. Для начала необходимо подготовить обучающий набор данных, где каждый клиент помечен как "отток" или "нет оттока", а также содержащий релевантные характеристики клиентов (например, длительность пользования услугами, количество звонков в службу поддержки, средний счет).

Задача 3: Рекомендательная система

Условие: Онлайн-магазин хочет создать рекомендательную систему, предлагающую клиентам товары, которые им могут понравиться. Какие методы можно использовать?

Ответ: Для создания рекомендательной системы можно использовать методы коллаборативной фильтрации (рекомендации на основе предпочтений других пользователей с похожими вкусами) и/или контент-based фильтрации (рекомендации на основе характеристик товара). Алгоритмы коллаборативной фильтрации включают user-based и item-based подходы. Контент-based фильтрация использует информацию о товарах (жанр, автор, описание).

Задача 4: Анализ текстовых данных

Условие: Компания хочет проанализировать отзывы клиентов о своем продукте, извлекая из них общие темы и настроения. Какие методы можно использовать?

Ответ: Для анализа текстовых данных можно использовать методы обработки

естественного языка (NLP). Это включает токенизацию (разбиение текста на слова), лемматизацию/стемминг (приведение слов к их основным формам), извлечение сущностей, анализ настроений (определение положительных, отрицательных или нейтральных отзывов). Для темы используются методы тематического моделирования (например, LDA).

Задача 5: Выявление аномалий

Условие: Компания заметила необычное поведение в своей системе платежей. Как обнаружить аномалии?

Ответ: Для обнаружения аномалий можно использовать методы анализа выбросов. Это включает статистические методы (например, определение выбросов по стандартному отклонению), методы на основе плотности (например, LOF – Local Outlier Factor), а также нейронные сети и машинное обучение.

Задача 6: Предсказание продаж

Условие: Магазин хочет предсказать свои продажи на следующий месяц. Какие методы подходят?

Ответ: Для предсказания продаж можно использовать методы регрессии, такие как линейная регрессия, регрессия с временными рядами (ARIMA, Prophet), нейронные сети (RNN). Необходимо подготовить исторические данные о продажах, возможно, учитывая внешние факторы (например, праздники, сезонность).

Задача 7: Оценка качества модели

Условие: Разработана модель классификации, предсказывающая удовлетворенность клиентов. Как оценить ее качество?

Ответ: Для оценки качества модели классификации можно использовать метрики, такие как точность (precision), полнота (recall), F1-мера, AUC-ROC (Area Under the ROC Curve). Также необходимо использовать кросс-валидацию для получения более надежной оценки обобщающей способности модели.

Задача 8: Выбор метода для задачи

Условие: Перед вами стоит задача сегментации клиентов по их поведению на сайте (просмотренные страницы, время, проведенное на сайте, покупки). Какой метод интеллектуального анализа данных вы бы выбрали?

Ответ: Для сегментации клиентов по их поведению на сайте наилучшим образом подходит метод кластеризации. В данном случае можно использовать k-means или иерархическую кластеризацию в зависимости от размера набора данных и желаемой

структуры кластеров. Предварительная обработка данных (нормализация, обработка пропущенных значений) является важным этапом