

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Ястребов Олег Александрович
Должность: Ректор
Дата подписания: 16.10.2025 16:38:25
Уникальный программный ключ:
sa953a0120d891083f959675078ef1a989dae18a

**Федеральное государственное автономное образовательное учреждение
высшего образования «Российский университет дружбы народов имени Патриса
Лумумбы»**

Высшая школа управления

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ И БАЛЛЬНО-РЕЙТИНГОВАЯ СИСТЕМА
ОЦЕНИВАНИЯ УРОВНЯ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ ПО
ДИСЦИПЛИНЕ**

Наука о данных для бизнеса

(наименование дисциплины/модуля)

**Оценочные материалы рекомендованы МССН для направления подготовки/
специальности:**

38.04.05 Бизнес-информатика

(код и наименование направления подготовки/специальности)

**Освоение дисциплины ведется в рамках реализации основной профессиональной
образовательной программы высшего образования (ОП ВО):**

Е-commerce-электронная коммерция

(наименование (профиль/специализация) ОП ВО)

Оценочные материалы разработаны/актуализированы для учебного года:

2026/2027

(учебный год)

Москва

1. БАЛЛЬНО-РЕЙТИНГОВАЯ СИСТЕМА ОЦЕНИВАНИЯ УРОВНЯ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ ПО ДИСЦИПЛИНЕ

Оценивание уровня сформированности компетенций по итогам изучения дисциплины «Наука о данных для бизнеса» осуществляется в соответствии с действующей в РУДН Балльно-рейтинговой системой (БРС).

Таблица 1.1. Балльно-рейтинговая система оценивания уровня сформированности компетенций по дисциплине/практике

Код контролируемой компетенции или ее части	Контролируемая тема дисциплины	ФОСы (формы контроля уровня освоения ООП)									Баллы темы
		Аудиторная работа					Самостоятельная работа		Промежуточная аттестация		
		Устный/ письменный	Тест	Коллоквиум	Контрольная работа	Коммуникация	Доклад	Реферат	Защита курсового проекта/ работы	Экзамен/ зачет	
УК-1 УК-2 ОПК-1 ОПК -3	Понятие и сущность науки о данных (Data Science). Жизненный цикл проекта по анализу данных: CRISP-DM, KDD, SEMMA методологии.	2				2					4

	Роль Data Science в цифровой трансформации бизнеса. Специфика данных в электронной коммерции: веб-аналитика, клиентские данные, товарные каталоги.										
УК-1 УК-2 ОПК-1 ОПК -3	Большие данные (Big Data) в e-commerce: концепция 5V (Volume, Velocity, Variety, Veracity, Value). Источники данных в электронной коммерции: логи веб-серверов, транзакционные данные, социальные сети, мобильные приложения. Инфраструктура для работы с большими данными: Hadoop, Spark, NoSQL базы данных.	2		5							7

УК-1 УК-2 ОПК-1 ОПК -3	Основы машинного обучения в бизнес-контексте. Типы задач машинного обучения: обучение с учителем, без учителя, с подкреплением. Метрики качества моделей машинного обучения. Валидация моделей: кросс-валидация, разделение на train/validation/test выборки. Борьба с переобучением и недообучением.	2			5				15		22
УК-1 УК-2 ОПК-1 ОПК -3	Алгоритмы классификации для e-commerce: логистическая регрессия, деревья решений, случайный лес, SVM, наивный байесовский классификатор. Практические применения:	2	5			2	5				14

	классификация клиентов, определение категорий товаров, fraud detection. Оценка качества классификации: accuracy, precision, recall, F1-score, AUC-ROC.										
УК-1 УК-2 ОПК-1 ОПК -3	Рекомендательные системы в e-commerce. Коллаборативная фильтрация: user-based и item-based подходы. Контентная фильтрация и гибридные рекомендательные системы. Матричная факторизация и алгоритмы глубокого обучения для рекомендаций. Evaluation метрики для рекомендательных систем: MAE, RMSE,	2				2		5			9

	Precision@K, Recall@K.										
УК-1 УК-2 ОПК-1 ОПК -3	Веб-аналитика и анализ поведения пользователей. Воронки конверсии и анализ пользовательских путей. А/В тестирование и экспериментальный дизайн в e-commerce. Статистическая значимость результатов экспериментов. Многовариантное тестирование и байесовские методы в А/В тестах.		4		5		5				14
УК-1 УК-2 ОПК-1 ОПК -3	Экзамен									30	30
Итого		10	9	5	10	6	10	5	15	30	100

2. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ И САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ ПО ДИСЦИПЛИНЕ

Соответствие систем оценок
(используемых ранее оценок итоговой академической успеваемости, оценок ECTS и балльно-рейтинговой системы оценок текущей успеваемости)

Оценка	Неудовлетворит.		Удовлетворительно	Хорошо		Отлично	
Оценка ECTS	F(2)	FX (2+)	E(3)	D(3+)	C (4)	B (5)	A (5+)
Максимальная сумма баллов							
100	менее 31	31-50	51-60	61-68	69-85	86-94	95-100

Описание оценок ECTS:

А ("Отлично") - теоретическое содержание курса освоено полностью, без пробелов, необходимые практические навыки работы с освоенным материалом сформированы, все предусмотренные программой обучения учебные задания выполнены, качество их выполнения оценено числом баллов, близким к максимальному.

В ("Очень хорошо") - теоретическое содержание курса освоено полностью, без пробелов, необходимые практические навыки работы с освоенным материалом в основном сформированы, все предусмотренные программой обучения учебные задания выполнены, качество выполнения большинства из них оценено числом баллов, близким к максимальному.

С ("Хорошо") - теоретическое содержание курса освоено полностью, без пробелов некоторые практические навыки работы с освоенным материалом сформированы недостаточно, все предусмотренные программой обучения учебные задания выполнены, качество выполнения ни одного из них не оценено минимальным числом баллов, некоторые виды заданий выполнены с ошибками.

Д ("Удовлетворительно") - теоретическое содержание курса освоено частично, но пробелы не носят существенного характера, необходимые практические навыки работы с освоенным материалом в основном сформированы, большинство предусмотренных программой обучения учебных заданий выполнено, некоторые из выполненных заданий, возможно, содержат ошибки.

Е ("Посредственно") - теоретическое содержание курса освоено частично, некоторые практические навыки работы не сформированы, многие предусмотренные программой обучения учебные задания не выполнены, либо качество выполнения некоторых из них оценено числом баллов, близким к минимальному.

FX ("Условно неудовлетворительно") - теоретическое содержание курса освоено частично, необходимые практические навыки работы не сформированы, большинство предусмотренных программой обучения учебных заданий не выполнено либо качество их выполнения оценено числом баллов, близким к минимальному; при дополнительной самостоятельной работе над материалом курса возможно повышение качества выполнения учебных заданий

Ф ("Безусловно неудовлетворительно") - теоретическое содержание курса не освоено, необходимые практические навыки работы не сформированы, все выполненные

учебные задания содержат грубые ошибки, дополнительная самостоятельная работа над материалом курса не приведет к какому-либо значимому повышению качества выполнения учебных заданий.

Реализация курса предусматривает интерактивные лекции, практические занятия (семинары) с использованием мультимедийного оборудования и интерактивного учебника, подготовку самостоятельных творческих работ и их последующие презентации, тестирование, проведение групповых дискуссий по тематике курса, современные технологии контроля знаний.

Темы рефератов.

1. Роль продуктовой аналитики и A/B-тестирования в росте юнит-экономики цифровых сервисов.
2. Качественные и количественные методы при формулировке гипотез в data-driven продакт-менеджменте.
3. Архитектура данных для бизнеса: дата-март vs. озеро данных и их влияние на скорость принятия решений.
4. Причинно-следственный вывод в бизнесе: разница-в-разницах, инструментальные переменные, матчинги.
5. Интерпретируемость моделей (SHAP, LIME) и доверие бизнеса к ML-решениям.
6. Персонализация предложений: рекомендательные системы в e-commerce и маркетплейсах.
7. Прогнозирование спроса: классические временные ряды vs. градиентный бустинг vs. нейросети.
8. Кредитный скоринг: баланс точности, интерпретируемости и регуляторных требований.
9. Антифрод в финтехе и e-commerce: поведенческие признаки, графовые методы, real-time scoring.
10. Маркетинговая атрибуция: last-click, data-driven, MMM и их сопоставление.
11. Ценообразование на основе данных: эластичность спроса и динамическое ценообразование.
12. Customer Lifetime Value (CLV): методы оценки и применение в сегментации/ретеншоне.
13. Управление оттоком (churn): раннее предупреждение, триггерные коммуникации, uplift-моделирование.
14. AIOps и MLOps: мониторинг, drift, retraining и экономический эффект устойчивых пайплайнов.
15. Генеративный ИИ для бизнеса: сценарии применения, риски галлюцинаций, оценка ROI.
16. Data Governance и качество данных: метрики, SLA и влияние на метрики бизнеса.
17. Этические и правовые аспекты работы с данными: bias, fairness, приватность, согласия.
18. Графовые базы и анализ сетей: кейсы в логистике, телекомах и маркетинге влияния.
19. Causal ML для оптимизации маркетинговых кампаний и персонализированных воздействий.
20. Экономика экспериментов: дизайн, статвлассть, минимально детектируемый эффект и риски бизнеса.

Критерии оценки

Реферат – самостоятельная работа студента, представляющая собой краткое изложение полученных результатов теоретического анализа определенной учебно-исследовательской темы, где автор раскрывает суть исследуемой проблемы, приводит различные точки зрения, а также собственные взгляды на нее. При оценке реферата необходимо учитывать следующие положения:

1. В тексте должно быть продемонстрировано владение предметом исследования, его понятийным аппаратом, терминологией, понимание современных тенденций и проблем в исследовании предмета.
2. Текст должен быть завершенным и четко структурированным, посвященным строго заданной проблематике.
3. Текст должен заканчиваться обоснованными выводами, полученными на основе анализа фактов и решающими поставленную в начале работы задачу.
4. Грамотность и логика изложения, терминология и лексика должны соответствовать заданной тематике и поставленной автором задаче.
5. Наличие правильно оформленных иллюстраций, ссылок на источники литературы.
6. Объем – не более 10 страниц А4 машинописного текста.

Оценивание: максимум 5 баллов, где 3 балла – содержание работы, стиль, правильность выводов, 1 балла – оформление согласно требованиям, 1 балла – защита, включающая презентацию.

Тесты

1. Что является первым шагом в data-driven решении бизнес-задачи?
А) Выбор модели
В) Сбор как можно больше данных
С) Формулировка бизнес-гипотезы и целевой метрики
D) Тюнинг гиперпараметров
Правильный ответ: С
2. Какая метрика корректна для несбалансированной бинарной классификации с редким положительным классом?
А) Accuracy
В) Precision@k
С) MSE
D) MAE
Правильный ответ: В
3. Что лучше всего использовать для оценки влияния изменения цены на спрос при наличии панельных данных по магазинам и времени?
А) K-Means
В) Difference-in-Differences
С) PCA
D) DBSCAN
Правильный ответ: В
4. Какой риск чаще всего связан с одновременным запуском множества А/В-экспериментов без коррекции?
А) Переобучение
В) Мультиколлинеарность
С) Инфляция ошибки первого рода (мульти-тестинг)

- D) Недостаток памяти
Правильный ответ: C
5. Какая практика относится к MLOps?
A) Линейная регрессия
B) Мониторинг data/model drift в продакшене
C) t-SNE визуализация
D) Oversampling
Правильный ответ: B
6. Что характеризует uplift-моделирование?
A) Прогноз абсолютной вероятности отклика
B) Оценка причинного эффекта коммуникации на индивида
C) Снижение дисперсии оценок
D) Борьба с мультиколлинеарностью
Правильный ответ: B
7. Какая метрика уместна для регрессии выручки?
A) ROC-AUC
B) MAP@K
C) MAPE/SMAPE
D) F1
Правильный ответ: C
8. Что такое data leakage?
A) Утечка данных в интернет
B) Использование признаков из будущего/валидации в обучении
C) Недостаток данных
D) Шум в данных
Правильный ответ: B
9. Какой метод уменьшает размерность, сохраняя максимум дисперсии?
A) PCA
B) KNN
C) Naive Bayes
D) XGBoost
Правильный ответ: A
10. При оценке CLV, что критично для корректности?
A) Максимальная точность моментного прогноза чеков
B) Учет оттока/удержания и дисконтирование будущих потоков
C) Использование только средних значений
D) Игнорирование сезонности
Правильный ответ: B
11. Что отличает рекомендательные системы на коллаборативной фильтрации?
A) Используют только контент атрибутов
B) Используют поведенческие матрицы пользователь-товар
C) Не масштабируются
D) Применимы только для текста
Правильный ответ: B
12. Какая практика снижает риск p-hacking?
A) Пост-hoc выбор метрик
B) Предрегистрация гипотез и планов анализа
C) Увеличение числа метрик
D) Многократный подгон порогов
Правильный ответ: B
13. Что такое fairness в моделях?
A) Максимум точности на тесте

- В) Отсутствие корреляции признаков
С) Справедливость прогнозов между группами по защищенным признакам
D) Минимизация времени инференса
Правильный ответ: С
14. Какая валидация корректна для временных рядов?
A) Случайный k-fold
B) Leave-one-out
C) Rolling/expanding window
D) Bootstrap по наблюдениям
Правильный ответ: С
15. В маркетинговой атрибуции data-driven подход означает:
A) Назначение 100% конверсии последнему клику
B) Равномерное распределение по каналам
C) Оценку вкладов каналов на основе данных и моделей переходов
D) Игнорирование показываемости
Правильный ответ: С
16. Для борьбы с мультиколлинеарностью в регрессии подходит:
A) L1/L2-регуляризация
B) Увеличение learning rate
C) SMOTE
D) One-hot encoding
Правильный ответ: А
17. Что лучше описывает понятие data governance?
A) Набор визуализаций
B) Управление доступами, качеством, каталогом и политиками работы с данными
C) Ускорение вычислений
D) Снижение стоимости облака
Правильный ответ: В
18. Какой показатель чаще всего является целевым в управлении оттоком подписочного сервиса?
A) NPS только
B) Конверсия в установку приложения
C) Доля оттока (churn rate) и удержание (retention)
D) Время отклика сервера
Правильный ответ: С
19. Что означает «контрафактическая оценка» эффекта?
A) Оценка точности на тесте
B) Оценка результата в альтернативном сценарии «а что если»
C) Подгон гиперпараметров
D) Оценка рисков безопасности
Правильный ответ: В
20. Что критично для валидности А/В-теста?
A) Малая выборка
B) Случайная рандомизация и контроль внешних факторов
C) Подбор пользователей вручную
D) Анализ только победивших экспериментов
Правильный ответ: В

Критерии оценки: 1 балл за каждый вопрос теста

3. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ

Практическое задание

Кейс 1. E-commerce: прогнозирование оттока и персонализированные кампании

Ситуация: Маркетплейс наблюдает рост оттока продавцов и покупателей. Требуется снизить churn и повысить повторные покупки.

Данные:

- Транзакции за 12–24 месяца (user_id, date, amount, category, device, channel).
- Лог активности (сессии, просмотры, добавления в корзину, клики по пуш/почте).
- Справочники: сегменты, источники трафика, характеристики продавцов.

Задачи:

1. Определить бизнес-метрики: churn, retention, повторная покупка в 30/60/90 дней, ARPU, LTV.
2. Построить когортный анализ и выявить сезонность/эффекты каналов.
3. Сформировать признаки для модели оттока (RFM, давность визита, глубина просмотра, ценовая чувствительность, жалобы, возвраты).
4. Обучить модель классификации оттока (baseline: логистическая регрессия; сравнить с XGBoost/LightGBM).
5. Оценить метрики на несбалансированных данных: ROC-AUC, PR-AUC, Precision@k, Recall@k.
6. Провести feature importance/SHAP-анализ, выделить управляемые факторы риска.
7. Спроектировать эксперименты: кому, что и когда отправлять (скидка, контент, лимит).
8. Приоритизация по ожидаемому uplift и затратам; рассчитать прирост выручки и снижение оттока.

Результат: дашборд с метриками оттока, отчет по модели, план A/B-тестов и ожидаемый финансовый эффект.

Кейс 2. Ритейл-логистика: прогноз спроса и оптимизация запасов

Ситуация: Сеть офлайн+онлайн-ритейла испытывает out-of-stock и избыточные запасы.

Данные:

- Продажи по SKU-Store-Day за 2 года, промо-календарь, цены, остатки.
- Внешние факторы: погода, праздники, события, мобильность.

Задачи:

1. Выбрать целевые метрики: MAPE/WMAPE по SKU-Store, уровень out-of-stock, оборачиваемость, lost sales.
2. Провести EDA: тренды, сезонность, промо-эластичность, каннибализация SKU.
3. Сформировать признаки: лаги, скользящие агрегаты, флаги промо, календари, погода.
4. Бенчмарк: Naive/SeasonalNaive/ETS/Prophet vs. градиентный бустинг/DeepAR/N-BEATS (по ресурсам).
5. Валидация: rolling window split, backtesting; выбрать горизонт (7/14/28 дней).
6. Интеграция с политикой запасов (S,S) или (s,S), сервис-уровень, штрафы за дефицит/излишки.
7. Симуляция экономического эффекта: снижение OOS, высвобождение оборотного капитала, рост маржи.

Результат: модель прогноза по SKU-Store с отчётом по качеству, рекомендации по параметрам пополнения и финансовой оценкой эффекта.

Критерии оценки: 5 баллов – правильно решенное практическое задание.

РАЗРАБОТЧИКИ:

доцент

Должность, БУП

Подпись

Абуева М.М.

Фамилия И.О.

РУКОВОДИТЕЛЬ БУП:

Заведующий кафедрой
цифрового менеджмента

Наименование БУП

Подпись

Назюта С. В.

Фамилия И.О.

РУКОВОДИТЕЛЬ ОП ВО:

Зав.кафедрой прикладной
экономики

Должность, БУП

Подпись

Островская А.А.

Фамилия И.О.