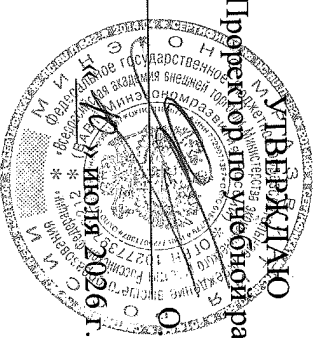


ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО
ОБРАЗОВАНИЯ «ВСЕРОССИЙСКАЯ АКАДЕМИЯ ВНЕШНЕЙ ТОРГОВЛИ
МИНИСТЕРСТВА ЭКОНОМИЧЕСКОГО РАЗВИТИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ»



О.М. Куликова

КОНТРОЛЬНО-ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

Б1.О.09 «Основы машинного обучения»

Уровень высшего образования

Магистратура

Направление подготовки

09.04.02 «Информационные системы и технологии»

Направленность (профиль)

«Информационные системы в экономике и управлении»

Квалификация – Магистр

Форма обучения – очная

Москва, 2026

Учебный план

Курс	Семестр	Трудоемкость, зет.	Академ. часы				Форма промежуточной аттестации
			Всего	Лекции	Практ занятия	Сам. работа	
1	1-2	6	216	24	24	168	Зачет, Диф. зачет

Проверяемые компетенции: ОПК-7

Максимальная итоговая сумма баллов, по которой может быть оценен уровень освоения изучаемой учебной дисциплины за семестр (далее – максимальная итоговая сумма баллов по дисциплине за семестр), не может превышать 100 баллов, которые включают оценку работы обучающегося в течение семестра и оценку, полученную на зачете / экзамене¹.

Оценка работы обучающегося в течение семестра составляет в сумме до 50 баллов. Структура оценки работы обучающегося в течение семестра включает отдельные баллы, начисляемые за:

- текущий контроль (выполнение тестов, контрольных работ, домашних заданий, подготовку презентаций, сдачу коллоквиумов, реферативный обзор дополнительного материала по теме и др.);
- работу на занятиях семинарского типа (выступления на занятиях, ответы на вопросы и др.);
- посещаемость занятий лекционного и семинарского типов.
- тестирование: количество баллов от количества правильных ответов.

Допускается проставление дополнительных «премиальных» баллов за отличную работу в семестре до 5 баллов. В таком случае, максимальное количество баллов за работу в семестре может быть увеличено до 55 баллов.

Критерии оценивания по мероприятиям текущего контроля и промежуточной аттестации прописываются подробно в рабочей программе учебной дисциплины.

Критерии оценки на зачете / экзамене/ дифференцированном зачете:

Оценка знаний, выявленных в ходе зачета / экзамена, составляет до 60 баллов.

Для выявления знаний по пропущенному в семестре материалу обучающемуся может быть предоставлена возможность ответить на дополнительные вопросы (задания), что может повысить итоговую оценку. За ответы на дополнительные вопросы (задания) допускается предоставление до 30 дополнительных баллов. В таком случае, максимальное количество баллов, которое обучающийся может получить в ходе зачета / экзамена, может быть увеличено до 90 баллов.

В случае, если итоговая сумма баллов по дисциплине (включающая оценку работы обучающегося, «премиальные» баллы за отличную работу в семестре, оценку за зачет / экзамен/ дифференцированный зачет и баллы за дополнительные вопросы (задания)) превышает 100 баллов, обучающемуся проставляется в ведомость 100 баллов.

¹Положение о балльно-рейтинговой системе оценки успеваемости обучающихся в федеральном государственном бюджетном образовательном учреждении высшего образования «Всероссийская академия внешней торговли Министерства экономического развития Российской Федерации», утвержденное приказом ВАВТ Минэкономразвития России от 30.06.2023г. № 293.

Перевод итоговой суммы баллов по учебной дисциплине из 100-балльной в эквивалент пятибалльной системы / системы зачета осуществляется по приведенным ниже шкалам.

Шкала перевода для зачета

Баллы по 100-балльной системе	Пятибалльная система оценки
52-100 баллов	Зачет
51 балл и ниже	Незачет
Шкала перевода для экзамена/дифференцированного зачета	
Баллы по 100-балльной системе	Пятибалльная система оценки
85-100 баллов	Отлично
70-84 баллов	Хорошо
52-69 баллов	Удовлетворительно
51 балл и ниже	Неудовлетворительно

Примеры тестовых заданий

№ п/п	Код компетенции	Задание	Ответ	Критерии оценивания
1.	ОПК-7.1	Какой метод оптимизации используется для обучения линейной регрессии? А) Метод Ньютона; Б) Градиентный спуск В) Симплекс-метод; Г) Метод наименьших квадратов	Г	1
2.	ОПК-7.1	Какая метрика используется для оценки качества регрессионных моделей? А) Точность; Б) F1-мера; В) Среднеквадратическая ошибка (MSE); Г) AUC-ROC	В	1
3.	ОПК-7.1	Раскройте математическую сущность градиентного спуска и назовите его основные модификации.	Градиентный спуск – итерационный метод минимизации функции потерь путём движения в направлении анти-градиента. Модификации: SGD (стохастический), мини-пакетный, Adam, RMSprop.	2
4.	ОПК-7.1	Какой метод регуляризации использует L1-штраф? А) Требуемая регрессия (Ridge); Б) LASSO В) ElasticNet; Г) Градиентный спуск	Б	1
5.	ОПК-7.1	Назовите метрики оценки качества регрессионных моделей и раскройте экономический смысл R^2 .	MAE, MSE, RMSE, MAPE, R^2 – коэффициент детерминации, доля	2

			дисперсии целевой переменной, объяснённая моделью (от 0 до 1).	
6.	ОПК-7.2	Какой метод относится к обучению с учителем? А) Кластеризация k-means; Б) Логистическая регрессия В) PCA; Г) DBSCAN	Б	1
7.	ОПК-7.2	Какая метрика используется для оценки качества классификационных моделей при дисбалансе классов? А) MAE; Б) F1-мера; В) R ² ; Г) MSE	Б	1
8.	ОПК-7.2	Назовите методы классификации данных и укажите условия их эффективного применения.	Логистическая регрессия (бинарная классификация); Метод опорных векторов (высокоразмерные данные); Решающие деревья (интерпретируемость); Случайный лес (высокая точность, устойчивость).	2
9.	ОПК-7.2	Какой метод кластеризации относится к плотностным методам? А) k-means; Б) DBSCAN; В) Иерархическая кластеризация; Г) PCA	Б	1
10.	ОПК-7.3	Для чего применяется метод главных компонент (PCA) в машинном обучении? А) Для классификации данных; Б) Для снижения размерности данных В) Для кластеризации; Г) Для построения рекомендаций	Б	1
11.	ОПК-7.3	Раскройте понятие «переобучение» в машинном обучении и назовите способы его предотвращения.	Переобучение – ситуация, когда модель слишком точно запоминает обучающую выборку, но плохо обобщается на новых данных. Способы предотвращения: регуляризация, уменьшение сложности модели, увеличение выборки, кросс-валидация, ранняя остановка.	2
12.	ОПК-7.3	Какой метод снижения размерности относится к нелинейным методам визуализации? А) PCA; Б) t-SNE; В) LDA; Г) SVD	Б	1
13.	ОПК-7.3	Что такое рекомендательные системы и какие типы рекомендаций существуют?	Рекомендательные системы – алгоритмы, предлагающие пользователям товары или контент на основе их предпочтений. Типы: коллаборатив-	2

			ная фильтрация, контентно-ориентированные рекомендации, гибридные системы.	
14.	ОПК-7.3	Какие метрики используются для оценки качества рекомендательных систем?	Точность (Precision), полнота (Recall), MAP (Mean Average Precision), NDCG (Normalized Discounted Cumulative Gain), покрытие и разнообразие рекомендаций.	2
15.	ОПК-7.1	Какой метод предобработки данных применяется для приведения признаков к единому масштабу? А) Нормализация; Б) Стандартизация; В) One-Hot Encoding; Г) Label Encoding	Б	1
16.	ОПК-7.2	Раскройте принцип работы метода k-ближайших соседей (k-NN) для классификации и регрессии.	Для классификации – голосование ближайших соседей (мажоритарное правило). Для регрессии – усреднение значений целевой переменной ближайших соседей. Выбор k влияет на сглаживание и чувствительность к шуму.	2
17.	ОПК-7.3	Что такое проблема холодного старта в рекомендательных системах и как она решается?	Проблема холодного старта – невозможность давать рекомендации новым пользователям или объектам без истории взаимодействий. Решения: использование контентно-ориентированных подходов, опросов, данных из профиля, демографической информации.	2
18.	ОПК-7.1	Какие этапы включает процесс построения модели машинного обучения?	1) Постановка задачи; 2) Сбор и предобработка данных; 3) Разделение выборки; 4) Выбор модели; 5) Обучение; 6) Оценка качества; 7) Настройка гиперпараметров; 8) Интерпретация и внедрение.	2
19.	ОПК-7.2	Раскройте принцип работы метода опорных векторов (SVM) для классификации.	SVM строит гиперплоскость, максимизирующую отступ (margin) между классами. Использует ядерный трюк для перевода данных в пространство	2

		наибольшей размерности для линейной разделённости в нелинейных случаях.	
--	--	---	--

Вопросы к зачёту (1 семестр)

1. Синтаксические конструкции языка Python и встроенные типы данных.
2. Структуры данных Python: списки, кортежи, словари, множества — особенности и области применения.
3. Библиотека NumPy: многомерные массивы, векторные и матричные операции.
4. Библиотека Pandas: структуры Series и DataFrame, операции чтения, записи, фильтрации и агрегации данных.
5. Библиотека Matplotlib: основы визуализации данных, построение графиков.
6. Библиотека SciPy: научные вычисления, статистические функции и оптимизация.
7. Определение машинного обучения как научной дисциплины. Классификация задач МО.
8. Основные понятия МО: объект, признак, целевая переменная, выборка, обучающий и тестовый наборы.
9. Типы признаков: количественные, порядковые, категориальные, бинарные — методы кодирования.
10. Понятие обобщающей способности модели, переобучения и недообучения.
11. Роль функций потерь и метрик качества в процессе обучения.
12. Постановка задачи регрессии: прогнозирование непрерывной целевой переменной.
13. Метод k-ближайших соседей (k-NN) для регрессии: принцип работы, выбор метрики и параметра k.
14. Линейная регрессия: математическая модель и метод наименьших квадратов.
15. Градиентный спуск как итерационный метод оптимизации: полный, стохастический и мини-пакетный варианты.
16. Выбор шага обучения, скорость сходимости и критерии останова в градиентном спуске.
17. Мультиколлинеарность и методы регуляризации: гребневая регрессия (L2), LASSO (L1) и ElasticNet.
18. Оценка качества регрессионных моделей: MAE, MSE, RMSE, R².
19. Постановка задачи бинарной и многоклассовой классификации.
20. Линейный классификатор: разделяющая гиперплоскость и уравнение линейного решающего правила.
21. Логистическая регрессия: сигмоидальная функция активации и интерпретация вероятностных оценок.
22. Метод опорных векторов (SVM) для линейно разделимых данных: максимизация отступа.
23. Метод k-ближайших соседей (k-NN) для классификации: принцип голосования и выбор параметра k.
24. Оценка качества классификационных моделей: матрица ошибок, точность, полнота, F1-мера.
25. ROC-кривая и AUC как метрики качества бинарной классификации.

ВОПРОСЫ К ДИФФЕРЕНЦИРОВАННОМУ ЗАЧЁТУ

1. Инструментальная среда Python: синтаксис, типы данных и структуры данных.
2. Библиотеки NumPy, Pandas, Matplotlib, SciPy: назначение и основные возможности.
3. Классификация задач машинного обучения: обучение с учителем, без учителя, с подкреплением.
4. Основные понятия МО: признаки, целевая переменная, выборка, переобучение, недообучение.
5. Типы признаков и методы их преобразования и кодирования.
6. Постановка задачи регрессии. Метод k-ближайших соседей для регрессии.

7. Линейная регрессия: модель, метод наименьших квадратов, нормальное уравнение.
8. Градиентный спуск и его модификации (SGD, Adam).
9. Регуляризация в линейной регрессии: Ridge, Lasso, ElasticNet.
10. Метрики оценки качества регрессионных моделей: MAE, MSE, RMSE, R².
11. Постановка задачи классификации. Линейный классификатор и разделяющая гиперплоскость.
12. Логистическая регрессия: сигмоида, функция потерь и интерпретация вероятностей.
13. Метод опорных векторов: линейный и нелинейный случаи, ядерный трюк.
14. Метрики оценки качества классификации: матрица ошибок, точность, полнота, F1-мера, ROC-AUC.
15. Методы сбора данных: веб-скрапинг, парсинг, использование API.
16. Этапы очистки и предобработки данных: пропуски, выбросы, нормализация.
17. Кодирование категориальных признаков: Label Encoding, One-Hot Encoding, Target Encoding.
18. Разделение выборки, кросс-валидация и её стратегии.
19. Структура решающего дерева: узлы, листья, правила ветвления.
20. Критерии качества разбиения: индекс Джини, энтропия, дисперсия.
21. Жадный алгоритм построения дерева и методы борьбы с переобучением (pruning).
22. Регрессионное дерево: особенности построения и обработка выбросов.
23. Случайный лес: бутстрэп, случайный выбор признаков, агрегирование.
24. Интерпретируемость решающих деревьев и важность признаков.
25. Постановка задач кластеризации: обнаружение скрытых структур в данных.
26. Метод k-means: алгоритм, выбор числа кластеров (метод локтя, силуэт).
27. Метод DBSCAN: плотностная кластеризация и преимущества.
28. Иерархическая кластеризация: агломеративные методы и дендрограмма.
29. Методы снижения размерности: PCA — математические основы и интерпретация.
30. Метод t-SNE: нелинейное вложение, особенности применения.
31. Понятие и назначение рекомендательных систем.
32. Коллаборативная фильтрация: user-based и item-based подходы.
33. Контентно-ориентированные рекомендательные системы.
34. Персонализированные и неперсонализированные рекомендации.
35. Метрики оценки качества рекомендательных систем: точность, полнота, MAP, NDCG.
36. Проблема холодного старта и способы её преодоления.
37. Современные подходы к построению рекомендательных систем на основе глубокого обучения.