Федеральное государственное бюджетное учреждение науки
Институт машиноведения им. А.А. Благонравова
Российской академии наук
(ИМАШ РАН)

УТВЕРЖДАЮ
Заместитель директора по вопросам
подготовки кадров высшей
квалификации ИМАШ РАН
_____ к.ф-м.н. Жданов Э.Р.
«_01_» _сентября_ 2026 года
м.п.

Комплект оценочных материалов
по дисциплине «Методы машинного обучения»

Направление подготовки:	27.04.07 Наукоёмкие технологии и экономика инноваций
Профиль подготовки:	Инновационное проектирование и разработка технологий
Квалификация:	Магистр
Форма обучения	Очная

Москва

Содержание

1. Общие положения....3
2. Результаты освоения дисциплины, подлежащие проверке....3
3. Контроль и оценка освоения учебной дисциплины....3
4. Оценочные материалы для проведения текущего контроля успеваемости....4
5. Оценочные материалы для проведения промежуточной аттестации....5

1. Общие положения

Комплект оценочных материалов (далее - ОМ) предназначен для контроля и оценки образовательных достижений обучающихся, освоивших программу дисциплины «Методы машинного обучения».

ОМ включают контрольные материалы для проведения текущего контроля успеваемости и промежуточной аттестации в форме зачета.

ОМ разработаны на основании:

- основной профессиональной образовательной программы по направлению подготовки 27.04.07 Наукоёмкие технологии и экономика инноваций;
- рабочей программы дисциплины «Методы машинного обучения».

2. Результаты освоения дисциплины, подлежащие проверке

В результате освоения дисциплины обучающийся должен обладать следующими компетенциями

ПК-2	**Способен организовывать и координировать сбор, обработку и анализ информации по результатам деятельности машиностроительной организации**
ПК-2.2	Способен формировать отчетность по результатам контроля и выявлять причины возникших отклонений

Код индикатора компетенции	Планируемые результаты обучения по дисциплине
ПК-2.2	*Знать:* - базовые алгоритмы машинного обучения (линейная/логистическая регрессия, деревья решений, K-Means) и их применимость для задач контроля; -метрики оценки качества моделей и методы их интерпретации (важность признаков). *Уметь:* - применять библиотеку scikit-learn для построения моделей прогнозирования и классификации; -интерпретировать результаты моделирования для выявления причин отклонений фактических данных от плановых. *Иметь практический опыт:* - разработки модели машинного обучения для прогнозирования KPI и выявления факторов, влияющих на них; - создания прототипа автоматизированной системы формирования отчетности с анализом причин отклонений.

3. Контроль и оценка освоения учебной дисциплины

Элемент учебной дисциплины	Текущий контроль		Промежуточная аттестация	
	Форма контроля	Проверяемые идк	Форма контроля	Проверяемые идк
Тема 1. Основы машинного обучения и задачи прогнозирования	Отчет по лабораторному практикуму	ПК-2.2		
Тема 2. Ансамблевые методы и деревья решений	Отчет по ситуационному практикуму	ПК-2.2		
Тема 3. Обучение без учителя: кластеризация и поиск аномалий	Отчет по лабораторному практикуму	ПК-2.2		
Тема 4. Построение	Отчет по	ПК-2.2		

сквозного ML-решения для отчетности	ситуационному практикуму			
			Зачет	ПК-2.2

4. Оценочные материалы для проведения текущего контроля успеваемости

Показатели и критерии оценивания компетенций на различных этапах их формирования, описание шкал оценивания

Ситуационный практикум
Отчет по практикуму: 25-21 – работа выполнена в срок, самостоятельно, использована требуемая информация, правильно сделаны необходимые выводы, хорошо аргументированы, даны исчерпывающие ответы на все поставленные вопросы; 20-15 – работа выполнена в срок, самостоятельно, использована требуемая информация, необходимые выводы сделаны частично, хорошо аргументированы, даны ответы на все поставленные вопросы; 14-9 – работа выполнена в срок, в основном самостоятельно, необходимые выводы сделаны частично, слабо аргументированы, даны ответы не на все вопросы; 8-1 – обучающийся подготовил работу несамостоятельно или не завершил в срок, выводы и ответы на вопросы отсутствуют.
Лабораторный практикум
25-21 – работа выполнена в срок, самостоятельно, правильно поняты и использованы соответствующие формулы, правильно определены соответствующие спецификации, использована требуемая информация, правильно выполнены требуемые расчеты, правильно выбраны совместимые комплектующие, сделаны необходимые выводы, хорошо аргументированы, даны исчерпывающие ответы на все поставленные вопросы; 20-15 – работа выполнена в срок, самостоятельно, правильно поняты и использованы соответствующие формулы, правильно определены соответствующие спецификации, использована требуемая информация, правильно выполнены требуемые расчеты, правильно выбраны совместимые комплектующие, необходимые выводы сделаны частично, хорошо аргументированы, даны ответы на все поставленные вопросы; 14-9 – работа выполнена в срок, в основном самостоятельно, использованы соответствующие формулы; определены соответствующие спецификации, имеются ошибки в расчетах; выбраны совместимые комплектующие необходимые, выводы сделаны частично, слабо аргументированы, даны ответы не на все вопросы; 8-1 – обучающийся подготовил работу несамостоятельно или не завершил в срок, описание спецификации содержит незначительные ошибки, выводы и ответы на вопросы отсутствуют.

Порядок начисления баллов

Если студент набрал достаточное количество баллов ТКУ по дисциплине (более 70 баллов) в течение семестра, он может отказаться от сдачи зачета, и результат будет выставлен автоматически по его согласию.

Типовые контрольные задания и иные материалы, необходимые для процедуры оценивания знаний, умений, навыков и(или) опыта деятельности

Типовые задания для лабораторных практикумов

Тема 1. Основы машинного обучения и задачи прогнозирования

«Прогнозирование себестоимости»

Студенту предоставляется чистый датасет с историческими данными о производстве (расход сырья, время работы станков, количество рабочих) и итоговой себестоимостью.

Что нужно сделать:

Реализовать полный цикл: загрузить данные, разделить на обучающую и тестовую выборки, обучить модель линейной регрессии, сделать прогноз и оценить его качество с

помощью метрики MAE (Mean Absolute Error).

Тема 3. Обучение без учителя: кластеризация и поиск аномалий

«Сегментация дефектов продукции»

Даны данные о браке: тип дефекта, время возникновения, станок, смена.

Что нужно сделать:

Применить алгоритм K-Means для автоматической группировки (кластеризации) этих данных. Определить оптимальное число кластеров. Проанализировать полученные группы и дать им интерпретацию (например, «Кластер 1 — дефекты из-за износа оборудования», «Кластер 2 — ошибки персонала»).

Типовые задания для ситуационных практикумов

Тема 2. Ансамблевые методы и деревья решений

Ситуация:

Компания столкнулась с тем, что проекты часто не укладываются в сроки. Нужно понять, почему. Данные содержат признаки (бюджет, количество участников, сложность, используемые технологии) и целевой признак (проект завершен в срок: да /нет).

Задача:

Построить модель (например, Random Forest), которая классифицирует проекты. Главная цель — не просто точность, а анализ важности признаков (Feature Importance), чтобы ответить на вопрос: «Какой фактор чаще всего приводит к срыву сроков?».

Тема 4. Построение сквозного ML-решения для отчетности

«Автоматизация отчета по контролю KPI».

Ситуация:

Руководителю нужен еженедельный отчет, который не просто показывает цифры, а объясняет их.

Задача:

Разработать скрипт, который берет новые данные о производстве, использует заранее обученную модель для прогноза KPI на следующую неделю, сравнивает прогноз с плановым значением и формирует вывод. Например: «Прогнозируемый объем производства ниже плана на 5%. Основная причина (по модели): снижение доступности ключевого станка №5».

5. Оценочные материалы для проведения промежуточной аттестации

Промежуточная аттестация по дисциплине проводится в форме зачета.

Зачет представляет собой выполнение обучающимся заданий билета, включающего в себя:

Задание №1 – теоретический вопрос на знание базовых понятий предметной области дисциплины, а также позволяющий оценить степень владения обучающегося принципами предметной области дисциплины, понимание их особенностей и взаимосвязи между ними.

Задание №2 – задание на анализ ситуации из предметной области дисциплины и выявление способности обучающегося выбирать и применять соответствующие принципы и методы решения практических проблем, близких к профессиональной деятельности.

Задание №3 – задания на проверку умений и навыков, полученных в результате освоения дисциплины.

Выполнение обучающимся заданий билета оценивается по следующей балльной шкале:

Задание 1: 0-30 баллов

Задание 2: 0-30 баллов

Задание 3: 0-40 баллов

«Зачтено»

– 90-100 – ответ правильный, логически выстроен, использована профессиональная терминология. Задания решены правильно. Обучающийся правильно интерпретирует полученный результат.

– 70 -89 – ответ в целом правильный, логически выстроен, использована профессиональная терминология. Ход решения заданий правильный, ответ неверный. Обучающийся в целом правильно интерпретирует полученный результат.

– 50 - 69 – ответ в основном правильный, логически выстроен, использована профессиональная терминология. Задание решено частично.

«Не зачтено»

– менее 50 – ответы на теоретическую часть неправильные или неполные. Задания не решены.

Типовые задания для проведения промежуточной аттестации обучающихся

Задания 1-го типа

1. История развития ИИ. Основные этапы.
2. Области искусственного интеллекта
3. Технологии работы с большими данными
4. Интеграция в бизнес-процессы. Направления, кейсы и условия применения. Методология управления проектами по анализу данных
5. Задача классификации. Постановка задачи, входные данные и алгоритм
6. Задача регрессии. Линейные модели. Переобучение
7. Метрики качества классификации
8. Метрики качества регрессии
9. Постановка задачи кластеризации. Основные алгоритмы кластеризации
10. Мягкая и жесткая кластеризация
11. Задача понижения размерности. Отбор признаков. Выделение признаков. Визуализация данных
12. Ассоциативные правила и рекомендательные системы
13. Задача обучения с подкреплением. Кумулятивная награда. Алгоритмы обучения с подкреплением
14. Ансамблевые методы: стекинг, бэггинг, бустинг. Решающие деревьяю Ансамблирование. Виды ансамблей
15. Введение в нейронные сети и история развития НС
16. Обучение нейронных сетей. Алгоритм обратного распространения ошибки
17. Нейронные сети в задачах машинного обучения
18. Популярные архитектуры нейронных сетей
19. Нейронные сети для анализа изображений. Задачи анализа изображений
20. Задача генерации изображений
21. Нейросети для работы с текстом. Анализ текстов. Векторные представления слов
22. Архитектуры нейросетей для анализа текстов
23. Задачи анализа и генерации текстов
24. Свёрточные нейронные сети. Введение. Операция свёртки. Свёрточный и пулинг слои. Техника Transfer Learning. Архитектуры
25. Рекуррентные нейронные сети. Введение. Forward pass. Backward pass. Примеры задач. Проблемы. Архитектура рекуррентного нейрона-LSTM. Схема LSTMнейрона-GPU. Двунаправленные рекуррентные нейронные сети
26. Статистический анализ выборочных данных. Графический анализ данных.
27. Формирование выборок и подготовка данных.
28. Статистические критерии. Критерии согласия.
29. Поиск взаимосвязей в данных и оценка их статистической значимости.
30. Одновыборочные и двухвыборочные критерии.
31. Сравнение нескольких выборок.

Задания 2-го типа

1. В чем разница задач классификации, регрессии, прогнозирования, кластеризации? Приведите примеры.
2. Какова взаимосвязь принципа максимума правдоподобия и принципа минимизации эмпирического риска?
3. В чем разница между ошибками I и II рода?
4. Какие три основных подхода к оцениванию плотности распределения вам знакомы?
5. В чем состоит метод редукции размерности Шурыгина?
6. Приведите основные шаги/блоки EM-алгоритма. Каков его критерий останова?
7. Опишите основную идею метода ближайших соседей. 8. Как производится отбор эталонных объектов с помощью алгоритмов СТОЛП/FRiSСТОЛП?
8. Какова природа связи линейного классификатора метода максимума правдоподобия?
9. Сформулируйте теорему Новикова о сходимости.
10. В чем состоит проблем переобучения?
11. Как решается задача настройки порога решающего правила по критерию числа ошибок I и II рода?
12. Опишите алгоритм обратного распространения ошибок для многослойных нейронных сетей.
13. Какие стратегии обучения для сетей Кохонена вам известны? Опишите их принципы.
14. Приведите несколько примеров задач кластеризации. Какие алгоритмы кластеризации вам известны?
15. Сформулируйте основные критерии выбора моделей.
16. Приведите основные положения комбинаторной теории обучения.
17. В чем основная сложность задачи отбора признаков? Какие алгоритмы могут применяться для решения этой задачи?
18. Опишите основные принципы работы алгоритма AdaBoost.
19. Какие различные определения информативности вам известны?
20. Опишите основные алгоритмы логическое классификации (достоинства, недостатки).
21. В чем состоит принцип частичной прецедентности?
22. Охарактеризуйте связь ассоциативного правила с понятием логической закономерности.
23. Приведите пример и формальную постановку задачи Semisupervised Learning?
24. Как адаптировать алгоритмы кластеризации для решения задач с частичным обучением?
25. В чем состоит задача коллаборативной фильтрации? Какие основные методы решения этой задачи?
26. Опишите метод максимума правдоподобия и метод максимума апостериорной вероятности для задачи тематического моделирования.
27. Сформулируйте задачу обучения с подкреплением. Какие основные алгоритмы решения этой задачи вам известны?

Задания 3-го типа

Задание № 1.

1). Создайте два вещественных тензора: a размером (3, 4) и b размером (12,)

2). Создайте тензор c, являющийся тензором b, но размера (2, 2, 3)

3). Выведите первый столбец матрицы a с помощью индексации

Задание №2.

1). Создайте два вещественных тензора: a размером (5, 2) и b размером (1,10)

2). Создайте тензор c, являющийся тензором b, но размера (5, 2) 3). Произведите все арифметические операции с тензорами a и c

Задание №3.

1). Создайте тензор целых чисел images размерности (100, 200, 200, 3) (можно интерпретировать это как 100 картинок размера 200x200 с тремя цветовыми каналами, то есть 100 цветных изображений 200x200). Заполните его нулями

2). Сделайте так, чтобы у i-ой по порядку картинки была нарисована белая полоса толщиной два пикселя в строках, которые соответствуют номеру картинки. Например, у 3-ей по порядку картинки белая полоска будет занимать 6 и 7 строки, у 99 -- 198 и 199. Сделать белую строку можно, записав в ячейки тензора число 255 (по всем трём каналам). Выведите пару примеров с помощью pylot

3). Посчитайте среднее тензора images по 1-ой оси (по сути - средняя картинка), умножьте полученный тензор на 70. Нарисуйте с помощью pyplot, должна получиться серая картинка (при взятии среднего нужно привести тензор к float с помощью .float(), при отрисовке обратно к int с помощью .int())

Задание №4.

Напишите функцию forward_pass(X, w) (w0 входит в w) для одного нейрона (с сигмоидой) с помощью PyTorch

Задание № 5.

1) Перейдите на Google Colaboratory, чтобы работать с GPU. Загрузите туда этот ноутбук (Upload notebook)

2) В Colab нажмите в меню сверху: Runtime -> Change Runtime Type -> GPU, где изначально стоит None

3) Объявите тензор a размера (2, 3, 4) и тензор b размера (1, 8, 3) на GPU, инцилизируйте их случайно равномерно (.uniform_())

4) Затем измените форму тензора b, чтобы она совпадала с формой тензора a, получите тензор c

5) Переместите c на CPU, переместите a на CPU

6) Оберните их в Variable()

7) Объявите тензор L = torch.mean((c - a)**2) и посчитайте градиент L по c (то есть ∂L∂c)