

ЗАДАНИЯ

заключительного этапа Универсиады РУДН
по предметному направлению
«Искусственный интеллект»

Уважаемый участник Универсиады RUDN-ON!

Перед завершающим этапом призываем обратить особое внимание на критерии оценки, представленные для данного этапа. Не менее важно также быть внимательным при выражении своих ответов. Ваша ясность мысли, точность и полнота ответов, соответствие заданным критериям, будут играть существенную роль в успешном завершении этапа. Желаем Вам сконцентрироваться, проявить все свои знания и умения наилучшим образом и не терять внимания к каждой детали. Удачи Вам в этом завершающем этапе Универсиады «RUDN-ON»!

Продолжительность – 120 минут.

Задание включает в себя 7 тестовых заданий закрытого типа, требующие выбора правильного ответа, а также 3 задачи, требующие решения.

По итогам заключительного этапа (финала) участник может набрать 100 баллов

Задания проверяют знания в области глубокого обучения, нейронных сетей, машинного обучения, оптимизации алгоритмов, обработки естественного языка или обработки изображения.

Критерии оценивания заданий заключительного этапа

Структура варианта финального этапа

Задание по направлению «Искусственный интеллект»
7 тестовых вопросов (по 7 баллов каждый, максимум 49 баллов)
3 задачи, требующие решения (единственное решение) (8 и 9 задача по 15 баллов, 10 задача – 21 балл)

Максимум 100 баллов

На рабочем месте участника не должно быть никаких посторонних предметов, за исключением:

- Компьютера или ноутбука, клавиатуры, мыши.
- Два-три листа бумаги формата А4 для записей, предоставленные оргкомитетом. Использование других бумажных носителей, например, тетрадей/блокнотов и др. запрещено
- Ручка (с чернилами черного или синего цвета);
- Вода в прозрачной емкости;
- Необходимые лекарства без упаковки;
- Оригинал документа, удостоверяющего личность.

Задания заключительного этапа Универсиады «RUDN-ON» по предметному направлению «Искусственный интеллект»

1. Какой класс функций наиболее широко используется в качестве функций активации в скрытых слоях современных глубоких нейронных сетей?

- (а) Сигмоидальные функции
- (б) Гиперболический тангенс
- (в) Кусочно-линейные функции (ReLU и её варианты)
- (г) Гауссовские функции

2. Трансформеры с механизмом внимания (Attention) наиболее эффективны для:

- (а) Обработки только изображений
- (б) Моделирования долгосрочных зависимостей в последовательностях
- (в) Исключительно табличных данных
- (г) Сжатия данных без потерь

3. Функция потерь _____

- (а) Измеряет разницу между предсказанными и фактическими значениями
- (б) Оптимизирует веса нейронной сети
- (в) Задаёт структуру нейронной сети
- (г) Контролирует скорость обучения нейронной сети

4. Какой алгоритм оптимизации является стандартом для обучения больших языковых моделей в 2025-2026 годах?

- (а) Стандартный градиентный спуск
- (б) AdamW с планировщиком скорости обучения
- (в) Метод Ньютона
- (г) Метод роя частиц

5. Переобучение нейронной сети возникает _____

- (а) Когда модель слишком хорошо запоминает обучающие данные, но плохо обобщает на новых
- (б) Когда нейронная сеть недостаточно хорошо соответствует обучающим данным
- (в) Когда используется слишком большой batch size
- (г) Только при использовании сверточных слоев

6. Архитектура Transformer состоит из 6 энкодерных блоков. Каждый блок содержит multi-head attention (8 голов) с размерностью модели $d_{\text{model}}=512$ и feed-forward сеть с промежуточным размером 2048.

Сколько приблизительно параметров содержит один энкодерный блок (без учета нормализации и dropout)?

- (a) ~2.1 млн
- (б) ~4.7 млн
- (в) ~8.3 млн
- (г) ~12.5 млн

7. Какая из перечисленных моделей НЕ является архитектурой для обработки естественного языка?

- (a) BERT
- (б) GPT
- (в) T5
- (г) ResNet

8. Известно, что x и y имеют линейную связь вида:

$$y = \alpha x$$

При этом коэффициент α определяется из N известных точек (x_i, y_i) с учетом L2-регуляризации для предотвращения переобучения.

Реализуйте программу, которая определит α , минимизирующей функцию потерь:

$$L(\alpha) = \sum (y_i - \alpha x_i)^2 + \lambda \alpha^2$$

где λ - коэффициент регуляризации.

Входные данные:

Первая строка: N (количество пар) и λ (коэффициент регуляризации), разделенные пробелом

Следующие N строк: пары целых чисел x и y , разделенных пробелом

Выходные данные:

Число α с точностью до второго знака после запятой.

9. Реализуйте программу для поиска минимума функции:

$$f(x) = 2x^4 - 3x^3 + 2x - 1$$

используя алгоритм оптимизации Adam (адаптивная оценка моментов).

Параметры:

$\alpha = 0.01$ (learning rate)

$\beta_1 = 0.9$ (момент первого порядка)

$\beta_2 = 0.999$ (момент второго порядка)

$\varepsilon = 10^{-8}$

Начальная точка: $x_0 = 0$

Максимум итераций: 10000

Критерий остановки: $|f'(x)| < 10^{-6}$

Выходные данные:

Точка минимума с точностью до одной десятой

10. Механизм внимания — ключевой компонент современных трансформеров и больших языковых моделей.

Реализуйте упрощенную версию **scaled dot-product attention**:

$$\text{Attention}(Q, K, V) = \text{softmax}(QK^T / \sqrt{d_k})V$$

Входные данные:

Первая строка: n (размер последовательности), d_k (размерность ключей)

Следующие n строк: векторы Q (query) — по d_k чисел в строке

Следующие n строк: векторы K (key) — по d_k чисел в строке

Следующие n строк: векторы V (value) — по d_k чисел в строке

Выходные данные:

n строк по d_k чисел — результат применения механизма внимания

Числа с точностью до 2 знаков после запятой

Примечание:

- $\text{Softmax}(x)_i = \exp(x_i) / \sum \exp(x_j)$
- Масштабирование на $\sqrt{d_k}$ предотвращает проблему затухающих градиентов.