

**Структура методических рекомендаций по подготовке к отборочному
этапу открытой универсиады федерального государственного автономного
образовательного учреждения высшего образования «Российский
университет дружбы народов имени Патриса Лумумбы» «RUDN-ON» в
2025/2026 уч. г.**

Интеллектуальные технологии и анализ данных в гуманитарной сфере

название предметного направления

Институт русского языка

1. О предметном направлении:

В рамках предметного направления «Интеллектуальные системы в гуманитарной среде» исследуются методология и практика проектирования, внедрения и критического осмысления технологий искусственного интеллекта и сложных информационных систем в контекст гуманитарных наук и человеко-ориентированных областей.

Данное направление представляет собой междисциплинарное исследовательское поле, основа которого заключается в применении инструментов цифрового анализа данных, геоинформационных систем, системного анализа и машинного обучения к обширным корпусам текстов, визуальных источников или культурных артефактов. Это позволяет выявлять ранее недоступные паттерны, закономерности и связи — будь то анализ эволюции литературных стилей, реконструкция исторических торговых путей или цифровое картографирование культурной памяти.

2. Отборочный этап:

Отборочный этап состоит из заданий по следующим дисциплинам:

1.1. Базовые дисциплины:

- Математический анализ
- Информатика и программирование

1.2. Специализированные дисциплины:

- Лингвистика
- Общее языкознание

- Линейная алгебра
- Теория вероятностей и математическая статистика
- Методы машинного обучения
- Базы данных

Количество вопросов: 25-30 вопросов

Продолжительность: 180 мин.

3. Список рекомендуемой литературы для подготовки:

1. Курош А.Г. Курс высшей алгебры. –М.: Наука, любой год издания.
2. Румшинский Л.З. Элементы теории вероятностей. –М.: Наука, 1966.
3. Сборник задач по математике для втузов. Под редакцией А.В.Ефимова и Б.П.Демидовича.–2-е изд., исправ. и доп.-Ч.1 – 2. –М.: Наука, 1986.
4. Задачи и упражнения по математическому анализу для втузов. Под редакцией Б.П.Демидовича. –М.: Наука, любой год издания.
5. Проскуряков И.В. Сборник задач по линейной алгебре. -М.: Наука, любой год издания.
6. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. –М.: Высшая школа, любой год издания.
7. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение. — 2-е изд., испр. — М.: ДМК Пресс, 2018 — 652 с.
8. Фостер Д. Генеративное глубокое обучение. Творческий потенциал нейронных сетей. — СПб.: Питер, 2020 — 336 с.
9. Гарсиа-Молина Г., Ульман Дж. Д., Уидом Дж. Системы баз данных. Полный курс. — М.: Издательский дом "Вильямс", 2003 — 1088 с.
10. Молинаро Э. SQL. Сборник рецептов. — СПб: Символ-Плюс, 2009. — 672 с.

ЗАДАНИЯ

заключительного этапа Универсиады РУДН
по предметному направлению
«Интеллектуальные технологии и анализ данных в
гуманитарной сфере»

Заключительный этап состоит из заданий по следующим дисциплинам:

1.1. Базовые дисциплины:

- Математический анализ
- Информатика и программирование

1.2. Специализированные дисциплины:

- ~ Линейная алгебра
- ~ Теория вероятностей и математическая статистика
- ~ Методы машинного обучения
- ~ Базы данных

Продолжительность: 120 мин.

Теория вероятностей и мат. статистика

1. Если вероятность успеха в одном испытании Бернулли равна p , то вероятность получить ровно k успехов в n испытаниях равна (3 балла):

- A) $C_k^n p^k (1-p)^{n-k}$
- B) $C_n^k p^k (1-p)^{n-k}$
- C) $\frac{n^k}{k!} p^k (1-p)^{n-k}$
- D) $\lambda^k \frac{e^{-\lambda}}{k!}$

2. Сформулируйте и докажите Локальную теорему Муавра-Лапласа (7 баллов)

3. Функция распределения $F(x)$ случайной величины X по определению равна (3 балла):

- A) Плотности вероятности в точке x
- B) $P(X < x)$
- C) Математическому ожиданию X
- D) $P(X = x)$

4. Распределение, описывающее число испытаний до первого успеха в схеме Бернулли, называется (3 балла):

- A) Биномиальным
- B) Пуассоновским
- C) Геометрическим
- D) Экспоненциальным

5. Математическое ожидание суммы двух случайных величин равно (3 балла)::

- A) Произведению их мат. ожиданий
- B) Сумме их мат. ожиданий
- C) Сумме их мат. ожиданий, только если они независимы
- D) Разности их мат. ожиданий

6. Дисперсия случайной величины $D[X] = 0$ означает, что (3 балла):

- A) Мат. ожидание равно 0
- B) Величина принимает только отрицательные значения
- C) Величина постоянна
- D) Величина имеет нормальное распределение

7. Коэффициент корреляции $\rho(X,Y)$ характеризует (3 балла):

- A) Причинно-следственную связь между X и Y
- B) Степень линейной стохастической зависимости между X и Y

- C) Независимость X и Y
- D) Разницу в дисперсиях X и Y

8. Закон больших чисел. Сформулируйте и докажите теорему. (7 баллов)

9. Центральная предельная теорема (ЦПТ) утверждает, что при выполнении условий (3 балла):

- A) Сумма большого числа независимых сл.в. сходится к нулю
- B) Распределение нормированной и центрированной суммы большого числа независимых сл.в. стремится к стандартному нормальному распределению
- C) Каждое слагаемое в сумме должно иметь нормальное распределение
- D) Дисперсии всех слагаемых должны быть равны 1

Алгебра и линейная алгебра

1. Аффинное пространство — это... (3 балла)
 - А) То же самое, что и векторное пространство
 - В) Множество точек с заданным началом координат
 - С) Множество точек, к которому присоединено векторное пространство, действующее как группа трансляций
 - D) Проективное пространство без бесконечно удалённых точек
2. Что такое кольцо вычетов по составному модулю m (3 балла)?
 - А) Всегда поле
 - В) Всегда область целостности
 - С) Кольцо с делителями нуля, если m — составное
 - D) Группа по сложению, но не по умножению
3. Коэффициент при a^2b^3 в разложении $(a+b)^5$ равен... (3 балла):
 - А) 5
 - В) 10
 - С) 15
 - D) 20
4. Доказать необходимое и достаточное условие линейной зависимости n -мерных векторов. (7 баллов)
5. Общее решение неоднородной СЛУ представляется в виде... (3 балла)
 - А) Линейной оболочки базисных векторов
 - В) Суммы частного решения неоднородной системы и общего решения соответствующей однородной системы
 - С) Множества всех линейно независимых решений
 - D) Фундаментальной системы решений
6. Сформулировать и доказать критерий положительной определённости квадратичной формы. (7 баллов)
7. Что такое подстановка (перестановка)? (3 балла)
 - А) Любое отображение множества в себя
 - В) Биективное отображение конечного множества на себя
 - С) Сюръективное отображение конечного множества на себя
 - D) Циклический сдвиг элементов множества
8. Правило Крамера применимо для решения СЛУ, если... (3 балла)
 - А) Система однородна
 - В) Число уравнений равно числу неизвестных
 - С) Определитель основной матрицы системы отличен от нуля

- D) Верны В и С

9. Поле вычетов существует по модулю m тогда и только тогда, когда... (3 балла)

- A) m — чётное число
- B) m — простое число
- C) m — степень двойки
- D) $m > 1$

10. Сформулировать и доказать теорему о ранге матрицы (7 баллов).

МАШИННОЕ ОБУЧЕНИЕ

1. Какое утверждение о «проблеме затухающего градиента» верно? (3 балла)

- А) При использовании функции ReLU градиент при обратном распространении может стать очень маленьким
- В) Без использования методов регуляризации, например Dropout, градиент с течением итераций обучения может стать очень маленьким.
- С) При использовании функции сигмоиды градиент при обратном распространении может стать очень маленьким
- D) При использовании слишком малой скорости обучения градиент может стать очень маленьким

2. Что минимизирует функция потерь «бинарная кросс-энтропия» в логистической регрессии? (3 балла)

- А) Сумму модулей отступов
- В) Логарифмическую ошибку на основе неверных предсказаний вероятностей
- С) Количество неверно классифицированных объектов
- D) Среднеквадратичное отклонение от эталонного значения

3. Есть обучающая выборка из двух примеров: (7 баллов)

Пример 1: вход $x_1=2$, истинный ответ $y_1=1$

Пример 2: вход $x_2=4$, истинный ответ $y_2=3$

Модель — это линейная регрессия с параметрами w и b , то есть имеет вид $y^{\wedge}=w \cdot x+b$.

Изначальные параметры: $w_0=0$, $b_0=0$.

Задача:

Используя среднеквадратическую ошибку (MSE), рассчитайте градиенты по параметрам w и b , чтобы сделать один шаг градиентного спуска с темпом обучения $\eta=0.1$.

Укажите подробное решение.

4. Есть модель логистической регрессии с весами $w=[0.5, -0.5]$, сдвигом $b=0$. Входные данные: $x=[3, 1]$ (7 баллов)

Истинный класс: $y=1$.

Значение функции активации для предсказания: $z = w_1 \cdot x_1 + w_2 \cdot x_2 + b$

Используя функцию $\sigma(z) = 1 + e^{-z}$

Рассчитайте градиенты функции потерь (кросс-энтропии) по весам и сдвигу, если использована одна итерация градиентного спуска с шагом $\eta = 0.1$. Ответ округлите до сотых.

5. Какому методу обучения соответствует следующее описание: модель обучается на данных без меток, и её задача — выявить структуру или закономерности в данных. Например, кластеризация или сокращение размерности. (3 балла)

- A) Обучение с подкреплением
- B) Обучение с учителем
- C) Обучение без учителя (ПРАВИЛЬНЫЙ)

Структура методических рекомендаций по подготовке к заключительному этапу открытой универсиады федерального государственного автономного образовательного учреждения высшего образования «Российский университет дружбы народов имени Патриса Лумумбы» «RUDN-ON» в 2025/2026 уч. г.

Интеллектуальные технологии и анализ данных в гуманитарной сфере

название предметного направления

Институт русского языка

1. О предметном направлении:

В рамках предметного направления «Интеллектуальные системы в гуманитарной среде» исследуются методология и практика проектирования, внедрения и критического осмысления технологий искусственного интеллекта и сложных информационных систем в контекст гуманитарных наук и человеко-ориентированных областей.

Данное направление представляет собой междисциплинарное исследовательское поле, основа которого заключается в применении инструментов цифрового анализа данных, геоинформационных систем, системного анализа и машинного обучения к обширным корпусам текстов, визуальных источников или культурных артефактов. Это позволяет выявлять ранее недоступные паттерны, закономерности и связи — будь то анализ эволюции литературных стилей, реконструкция исторических торговых путей или цифровое картографирование культурной памяти.

2. Заключительный этап:

Заключительный этап состоит из заданий по следующим дисциплинам:

1.1. Базовые дисциплины:

- Математический анализ
- Информатика и программирование

1.2. Специализированные дисциплины:

- Лингвистика
- Общее языкознание

- Линейная алгебра
- Теория вероятностей и математическая статистика
- Методы машинного обучения
- Базы данных

Количество вопросов: 25-30 вопросов

Продолжительность: 180 мин.

*Необходимо указать состав и тематику заданий отборочного этапа
(количество вопросов, количество баллов за вопрос)*

Необходимо указать продолжительность отборочного этапа

3. Список рекомендуемой литературы для подготовки:

1. Курош А.Г. Курс высшей алгебры. –М.: Наука, любой год издания.
2. Румшинский Л.З. Элементы теории вероятностей. –М.: Наука, 1966.
3. Сборник задач по математике для втузов. Под редакцией А.В.Ефимова и Б.П.Демидовича.–2-е изд., исправ. и доп.-Ч.1 – 2. –М.: Наука, 1986.
4. Задачи и упражнения по математическому анализу для втузов. Под редакцией Б.П.Демидовича. –М.: Наука, любой год издания.
5. Проскуряков И.В. Сборник задач по линейной алгебре. -М.: Наука, любой год издания.
6. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. –М.: Высшая школа, любой год издания.
7. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение. — 2-е изд., испр. — М.: ДМК Пресс, 2018 — 652 с.
8. Фостер Д. Генеративное глубокое обучение. Творческий потенциал нейронных сетей. — СПб.: Питер, 2020 — 336 с.
9. Гарсиа-Молина Г., Ульман Дж. Д., Уидом Дж. Системы баз данных. Полный курс. — М.: Издательский дом "Вильямс", 2003 — 1088 с.
10. Молинаро Э. SQL. Сборник рецептов. — СПб: Символ-Плюс, 2009. — 672 с.

4. Демоверсии заданий заключительного этапа (тесты):

1. Сколько существует биективных отображений (биекций) из множества из 4 элементов в себя?
 - А) 8
 - В) 16

C) 24 (Правильный. Это $4! = 24$)

D) 64

2. Какова вероятность вытянуть туза из хорошо перетасованной колоды в 36 карт? (Классическое определение)

A) $1/36$

B) $1/12$

C) $1/9$ (Правильный. 4 туза / 36 карт = $1/9$)

D) $1/6$

3. В архитектуре Transformer, что такое механизм внимания и как он отличается от рекуррентных сетей (RNN) в обработке последовательностей? Укажите ключевой компонент внимания.

A) Внимание — это механизм, который игнорирует порядок токенов, а RNN обрабатывают последовательности последовательно.

B) Внимание вычисляет веса для всех пар токенов одновременно с использованием матриц запросов, ключей и значений, позволяя параллельную обработку, в отличие от последовательной в RNN. (Правильный ответ: B)

C) Внимание использует только позиционные эмбединги, а RNN — только контекстные.

D) Внимание — это замена softmax в RNN, без изменений в параллелизме.