



Традиционно направление проводилось совместно с Райффайзен Банком, всегда желающим участникам удачи и веры в собственные силы. Помимо основных льгот 10 дипломантов направления получили подарки от Райфа и пропуск первых этапов отбора на любую из карьерных программ банка в 2022 году.

May the force be with you!

Вопрос **Инфо**

Уважаемые участники!

Олимпиадное задание по направлению «Прикладная математика и информатика» состоит из двух частей:

Инвариантная часть представлена заданиями № 1–6. Их нужно выполнить всем участникам.

Вариативная часть разделена на треки:

- Трек «Анализ данных и искусственный интеллект»: задания № 7 (раздел «Важно!»), 9, 10, 11.
- Трек «Современные компьютерные науки»: задания № 7–8 (раздел «Важно!»), 12, 13.
- Трек «Финансовые технологии»: задания № 8 (раздел «Важно!»), 14, 15, 16.

Вы можете сосредоточиться на выполнении заданий одного трека (чтобы претендовать на статус дипломанта I, II, III степени) или постараться решить наибольшее число задач вне зависимости от треков, чтобы претендовать на статус медалиста.

Во время выполнения заданий вы можете:

- пользоваться черновиком (в качестве черновика разрешено использовать чистые листы бумаги), но на проверку он не предъявляется;
- использовать встроенный в систему калькулятор;
- использовать таблицы математической статистики. Использование других справочных материалов и сторонних ресурсов строго запрещено.

[Нажмите, чтобы открыть справочные материалы](#)

Верим в ваш успех!

Вопрос 1

Балл: 10,00

Инвариантная часть. Это задание нужно выполнить всем участникам.

Исследуйте на сходимость ряд

$$\sum_{n=1}^{+\infty} x_n$$

где величина x_n задана рекуррентно:

$$x_{n+1} = x_n^3 - x_n, \quad x_1 = 1/2$$

Вопрос 2

Балл: 10,00

Инвариантная часть. Это задание нужно выполнить всем участникам.

Размерность образа линейного оператора $\varphi : \mathbb{R}^{17} \rightarrow \mathbb{R}^{17}$ равна 9. Какое максимальное и какое минимальное число различных вещественных собственных значений может иметь такой оператор?

Вопрос 3

Балл: 10,00

Инвариантная часть. Это задание нужно выполнить всем участникам.

Из урны с 50 белыми и 50 черными шарами случайно вытаскивают 10 (без возвращения). Какова вероятность того, что первыми были вытаснены 3 белых шара, при условии, что из 10 шаров ровно 3 оказались белыми, а все остальные – черными?

Вопрос 4

Балл: 10,00

Инвариантная часть. Это задание нужно выполнить всем участникам.

Сколько существует нестрогих линейных порядков, определенных на множестве $\{a, b, c, d, e, f, g, h, i, k\}$, таких, что

$$a \succ b \succ c \text{ и } d \succ e?$$

Вопрос 5

Балл: 10,00

Инвариантная часть. Это задание нужно выполнить всем участникам.

Найти все квадратные корни из матрицы A:

$$A = \begin{pmatrix} 7 & 9 & -12 \\ 10 & 16 & -22 \\ 8 & 12 & -16 \end{pmatrix},$$

то есть все такие матрицы B, что $B^2 = A$.

Вопрос **6**

Балл: 10,00

Инвариантная часть. Это задание нужно выполнить всем участникам.

Набор отрезков задан с помощью массива начал $B[1:n]$ и массива концов $E[1:n]$, в каждом из которых содержатся целые числа, причём гарантируется, что для любого $i: B[i] < E[i]$. Предложите алгоритм, для каждого i находящий индекс j такой, что отрезок $[B[j], E[j]]$ -- самый близкий к $[B[i], E[i]]$ из отрезков, которые лежат правее $[B[i], E[i]]$ и не пересекаются с ним, или (-1), если такого отрезка не существует. Степень близости определяется разностью $(B[j] - E[i])$. Если отрезков с одинаковой близостью несколько, достаточно найти какой-либо из них. Алгоритм должен работать за $O(n^2)$ операций (о-маленькое!).

Обоснуйте работоспособность вашего алгоритма. Оцените число операций, а также количество дополнительной памяти, которое будет использовать алгоритм.

В данном задании нельзя апеллировать к реализациям типов данных, структур данных и функций, доступных в тех или иных языках программирования. Все используемые алгоритмы должны быть подробно описаны.

Вопрос **Инфо**

Следующие два задания учитываются в рейтинге сразу по двум трекам.

Задание 7 учитывается с рейтингах по трекам «Анализ данных и искусственный интеллект» и «Современные компьютерные науки».

Задание 8 учитывается с рейтингах по трекам «Современные компьютерные науки» и «Финансовые технологии».

Вопрос **7**

Балл: 10,00

Это задание учитывается в рейтинге по трекам «Анализ данных и искусственный интеллект» и «Современные компьютерные науки», а также в рейтинге медалистов.

Неориентированный граф задан своей матрицей.

- Верно ли, что он планарен?
- Если он планарен, то верно ли, что при добавлении любого ребра (не добавляя вершин) граф становится не планарен? Если он не планарен, то верно ли, что если убрать любое ребро (не добавляя вершин) граф становится планарен?

Матрица смежности граф

а

	a	b	c	d	e	f	g	h
a	0	1	0	0	1	1	1	1
b	1	0	1	1	0	1	0	1
c	0	1	0	1	0	0	1	1
d	0	1	1	0	1	1	1	0
e	1	0	0	1	0	1	1	0
f	1	1	0	1	1	0	0	0
g	1	0	1	1	1	0	0	1
h	1	1	1	0	0	0	1	0

Вопрос 8

Балл: 10,00

Это задание учитывается в рейтинге по трекам Современные компьютерные науки и «Финансовые технологии», а также в рейтинге медалистов.

Два друга Илья и Дима решили сыграть в следующую игру. Каждый из них подбрасывает стандартную игральную кость, а потом они вместе складывают выпавшие числа. Назовем эту процедуру раундом игры. Если в первом раунде сумма выпавших чисел равна 7 или 11, то выигрывает Дима, а если она равна 2, 3 или 12, то выигрывает Илья. Если же сумма выпавших чисел в первом раунде оказалась равной числу N из множества $\{4, 5, 6, 8, 9, 10\}$, то игра продолжается. А именно, Илья и Дима проводят раунды до тех пор, пока в результате в каком-то из раундов не получится либо опять число N (из первого раунда) либо число 7. При этом, если получается число N , то выигрывает Дима. Какова вероятность выигрыша у Димы?

Вопрос 9

Балл: 10,00

Задание по треку «Анализ данных и искусственный интеллект». Это задание учитывается в рейтинге по треку «Анализ данных и искусственный интеллект» и в рейтинге медалистов.

Рассмотрим язык логики предикатов, содержащий один функциональный символ $+$ и один предикатный символ $=$ (оба двуместные), и рассмотрим интерпретации этого языка на абелевых группах $\mathbb{Z} \oplus \mathbb{Z}$ и $\mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z}$ (Элементы этих групп – соответственно, пары и тройки целых чисел, сложение – покомпонентное.) Существует ли формула логики предикатов (без свободных переменных), истинная в первой интерпретации и ложная во второй?

Вопрос 10

Балл: 10,00

Задание по треку «Анализ данных и искусственный интеллект». Это задание учитывается в рейтинге по треку «Анализ данных и искусственный интеллект» и в рейтинге медалистов.

Федеральное Дорожное Агенство совместно с Минстроем разрабатывает экспериментальную модель застройки кварталов для минимизации времени в пути от окраин до центра города. Предлагается осуществлять застройку кварталов шестиугольниками с кольцевыми развязками на стыке трёх кварталов. Вам необходимо оценить сколько районов смогут доехать до центра за N кварталов.

Для воспроизведения плана застройки нарисуйте три правильных шестиугольника, примыкающие друг к другу так что у каждого шестиугольника по одной общей грани с соседом. Будем считать грани шестиугольников ребрами графа степени 1, а углы – вершинами. Чтобы получить граф степени 2 (город расширяется), достройте шестиугольники так, чтобы все вершины, имеющие на предыдущем шаге три ребра были окружены тремя примыкающими друг к другу правильными шестиугольниками, имеющими одну общую грань с каждым соседом. Рассчитайте количество узлов, достижимых за N шагов из центра графа степени K построенному как описано выше при $N < K$ и объясните ответ.

Вопрос 11

Балл: 10,00

Задание по треку «Анализ данных и искусственный интеллект». Это задание учитывается в рейтинге по треку «Анализ данных и искусственный интеллект» и в рейтинге медалистов.

Нейросеть ничему не учится и каждую эпоху просто показывает случайные значения функции потерь: l_i --- последовательность независимых одинаково распределённых случайных величин из некоторого непрерывного распределения. Обучение останавливается в тот момент, когда некоторое значение функции потерь оказывается не лучше предыдущего: $l_i \geq l_{i-1}$. Сколько в среднем значений функции потерь будет посчитано (включая самое первое)?

Вопрос 12

Балл: 10,00

Задание по треку «Современные компьютерные науки». Это задание учитывается в рейтинге по треку «Современные компьютерные науки» и в рейтинге медалистов.

Пусть $\mathbf{X}^n = \{X_1, \dots, X_n\} \sim \text{Uniform}(0, \theta)$ --- выборка из равномерного распределения на отрезке $[0, \theta]$, где θ --- неизвестный параметр. Рассмотрим оценку этого параметра $\hat{\theta} = \max\{X_1, \dots, X_n\}$. Проверьте данную оценку на несмещённость и состоятельность.

Вопрос 13

Балл: 10,00

Задание по треку «Современные компьютерные науки». Это задание учитывается в рейтинге по треку «Современные компьютерные науки» и в рейтинге медалистов.

В задаче линейной регрессии $y \sim \langle x, w \rangle$ (здесь $\langle \cdot, \cdot \rangle$ означает скалярное произведение)

можно рассматривать различные функции потерь; в частности, такую:

$$\mathcal{L}(X, y, w) = \frac{1}{N} \sum_{i=1}^N \log(\cosh(\langle x_i, w \rangle - y_i)),$$

где

$$\cosh = \frac{1}{2}(e^x + e^{-x}), \text{ а } x_1, \dots, x_D \text{ --- столбцы матрицы } X.$$

Верно ли, что задача

$$\mathcal{L}(X, y, w) \rightarrow \min_w$$

имеет не более одного решения, если столбцы матрицы X линейно независимы? Обязательно объясните свой ответ. Какой смысл в множителе $1/N$ перед суммой, если он всё равно не влияет на поиск минимума? Сравните эту функцию потерь со среднеквадратичной

$$MSE(X, y, w) = \frac{1}{N} \sum_{i=1}^N (\langle x_i, w \rangle - y_i)^2$$

или с функцией потерь

$$MAE(X, y, w) = \frac{1}{N} \sum_{i=1}^N |\langle x_i, w \rangle - y_i|.$$

Как они меняются при росте по модулю отклонения $\langle x_i, w \rangle - y_i$? Что это означает для выучиваемых с их помощью моделей?

Вопрос 14

Балл: 10,00

Задание по треку «Финансовые технологии». Это задание учитывается в рейтинге по треку «Финансовые технологии» и в рейтинге медалистов.

Пусть $g(t)$ – количество газа (млн.куб.м) в газохранилище в момент времени t . Известно, что со скважины газ подается в газохранилище непрерывно со скоростью $g(t)/2$, а отбор газа из газохранилища происходит непрерывно со скоростью $g^2(t)/2$. Найдите функцию $g(t)$, если $g(0) = 100$.

Вопрос 15

Балл: 10,00

Задание по треку «Финансовые технологии». Это задание учитывается в рейтинге по треку «Финансовые технологии» и в рейтинге медалистов.

Вы рассматриваете задачу оценки линейной регрессии по 20 точкам, при этом предполагается, что условия теоремы Гаусса-Маркова выполнены:

$$\hat{y}_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t$$

Результаты МНК-оценок коэффициентов: $\hat{\beta}_0 = 2,4, \hat{\beta}_1 = 6,9, \hat{\beta}_2 = 5,1$.

Стандартные ошибки оценок коэффициентов соответственно:

$$\hat{\sigma}_0 = 0,6, \hat{\sigma}_1 = 0,3, \hat{\sigma}_2 = 9,8.$$

Кроме того, сумма квадратов $TSS = \sum_{t=1}^{20} (y_t - \bar{y})^2 = 2000$, а $RSS = \sum_{t=1}^{20} (y_t - \hat{y}_t)^2 = 200$.

- **A)** Проверьте на 1%-уровне значимости гипотезу о том, что коэффициент при переменной x_{1t} равен 7.

- **Б)** Предположим, что Вы добавили еще две переменные в модель x_{3t} и x_{4t} . Показатель R^2 в новой модели увеличился до 0,95. Какой тест позволяет оценить, насколько статистически значимо улучшилось качество модели? Проведите его и сделайте соответствующие выводы на уровне значимости 1%.

Вопрос **16**

Балл: 10,00

Задание по треку «Финансовые технологии». Это задание учитывается в рейтинге по треку «Финансовые технологии» и в рейтинге медалистов.

Рассмотрим модель бинарной классификации для случайной величины y , используя как регрессор скалярную случайную величину x :

$$Pr(y = 1|x) = F(\alpha + \beta x)$$

где $F(\bullet)$ - некоторая непрерывная функция распределения. Покажите, что в случае, если случайная величина x принимает только два различных значения, то значение функции максимального правдоподобия, рассчитанное в точке своего максимума (то есть при подстановке оценок максимального правдоподобия для α и β) не зависит от конкретного вида функции $F(\bullet)$.