

Задания заключительного этапа олимпиады школьников СПбГУ по математическому моделированию и искусственному интеллекту 9-11 классы

Вариант задания заключительного этапа Олимпиады школьников СПбГУ по математическому моделированию и искусственному интеллекту 2024/25 учебного года состоял из трёх задач разных типов в рамках тематики Олимпиады и 5 тестовых заданий на анализ данных. Каждая задача оценивалась определенным количеством баллов в зависимости от уровня сложности.

Наибольшая итоговая сумма баллов, которой оценивались ответы всех задач, была равна 100. Подсчет итоговой оценки осуществлялся путем суммирования баллов, выставленных за ответы каждой из задач. Оценка ответов по каждой задаче формировалась в зависимости от близости введённого ответа к точному ответу в соответствии с критериями оценки.

Задача 1: Моделирование и прогнозирование

Формулировка (Вариант 1)

Геннадий открыл свой небольшой онлайн-магазин. Начиная с января 2023 года он отмечал количество проданной продукции за неделю. Эти данные он собирал два года (104 недели), они находятся в прикрепленном файле “forecast1.csv”. Помогите Геннадию спрогнозировать объёмы продаж в течении первых 12-ти недель 2025-го года.

Изучая график объёма продаж, Геннадий решил, что наиболее подходящей моделью, которая описывает данные продаж, является функция

$$\text{model}(t) = a_1 + a_2 t + a_3 t^2 + a_4 \sin\left(2\pi \frac{1}{13} t\right)$$

где a_1, a_2, a_3, a_4 — это неизвестные параметры модели, а t — это индекс недели, начиная с 0. Найдите неизвестные параметры модели, при которых достигается минимальная квадратичная ошибка на имеющихся данных, то есть достигается минимум функции

$$\sum_{t=0}^{103} (\text{model}(t) - \text{sales}(t))^2$$

здесь $\text{sales}(t)$ — это значение объёма продаж за неделю с индексом t , эти значения берутся из файла. С помощью найденных оптимальных параметров модели сделайте предсказание значений объёмов продаж на следующие 12 недель после последней недели, отмеченной в файле.

В качестве ответа введите 12 полученных чисел через запятую.

Пример ввода: 1.01, 2.02, 3.03, 4.04, 5.05, 6.06, 7.07, 8.08, 9.09, 10.10, 11.11, 12.12

Формулировка (Вариант 2)

Геннадий открыл свой небольшой онлайн-магазин. Начиная с января 2023 года он отмечал количество проданной продукции за неделю. Эти данные он собирал два года (104 недели), они находятся в прикрепленном файле “forecast2.csv”. Помогите Геннадию спрогнозировать объёмы продаж в течении первых 12-ти недель 2025-го года.

Изучая график объёма продаж, Геннадий решил, что наиболее подходящей моделью, которая описывает данные продаж, является функция

$$\text{model}(t) = a_1 + a_2 t + a_3 t^2 + a_4 \sin\left(2\pi \frac{1}{13} t\right)$$

где a_1, a_2, a_3, a_4 — это неизвестные параметры модели, а t — это индекс недели, начиная с 0. Найдите неизвестные параметры модели, при которых достигается минимальная квадратичная ошибка на имеющихся данных, то есть достигается минимум функции

$$\sum_{t=0}^{103} (\text{model}(t) - \text{sales}(t))^2$$

здесь $\text{sales}(t)$ — это значение объёма продаж за неделю с индексом t , эти значения берутся из файла. С помощью найденных оптимальных параметров модели сделайте предсказание значений объёмов продаж на следующие 12 недель после последней недели, отмеченной в файле.

В качестве ответа введите 12 полученных чисел через запятую.

Пример ввода: 1.01, 2.02, 3.03, 4.04, 5.05, 6.06, 7.07, 8.08, 9.09. 10.10, 11.11, 12.12

Задача 2: Анализ аудиосигналов

Формулировка (Вариант 1)

Виктор приехал на школу анализа данных. Обучение проходило в командах по 11 человек.

Однако, свою команду требовалось найти, связавшись с ними по мобильному телефону. Чтобы выяснить номера, требовалось проанализировать аудио записи с набором этих мобильных номеров в тональном режиме. Помогите Виктору найти свою команду.

В прикреплённом ZIP-архиве “Var1.zip” находятся файлы в формате “.wav”, которые содержат аудио записи с набором в тональном режиме 10 мобильных номеров.

Для каждой цифры, набранной в тональном режиме, формируется звуковой сигнал, состоящий из двух синусоид вида

$$f(t) = \sin(2\pi\omega_1 t) + \sin(2\pi\omega_2 t)$$

где ω_1, ω_2 — это частоты. Каждой цифре соответствует уникальная пара частот, которая берётся из таблицы

	1209 Гц	1336 Гц	1477 Гц
697 Гц	1	2	3
770 Гц	4	5	6
852 Гц	7	8	9
941 Гц	*	0	#

Требуется по аудио записям расшифровать набранные мобильные номера. В качестве ответа введите 10 мобильных номеров. Пример ввода ответа:

89000000000
 89111111111
 89222222222
 89333333333
 89444444444
 89555555555
 89666666666
 89777777777
 89888888888
 89999999999

Формулировка (Вариант 2)

Виктор приехал на школу анализа данных. Обучение проходило в командах по 11 человек. Однако, свою команду требовалось найти, связавшись с ними по мобильному телефону. Чтобы выяснить номера, требовалось проанализировать аудио записи с набором этих мобильных номеров в тональном режиме. Помогите Виктору найти свою команду.

В прикрепленном ZIP-архиве “Var2.zip” находятся файлы в формате “.wav”, которые содержат аудио записи с набором в тональном режиме 10 мобильных номеров.

Для каждой цифры, набранной в тональном режиме, формируется звуковой сигнал, состоящий из двух синусоид вида

$$f(t) = \sin(2\pi\omega_1 t) + \sin(2\pi\omega_2 t)$$

где ω_1, ω_2 — это частоты. Каждой цифре соответствует уникальная пара частот, которая берётся из таблицы

	1209 Гц	1336 Гц	1477 Гц
697 Гц	1	2	3
770 Гц	4	5	6
852 Гц	7	8	9
941 Гц	*	0	#

Требуется по аудио записям расшифровать набранные мобильные номера. В качестве ответа введите 10 мобильных номеров. Пример ввода ответа:

89000000000
89111111111
89222222222
89333333333
89444444444
89555555555
89666666666
89777777777
89888888888
89999999999

Задача 3: Моделирование методом Монте-Карло

Формулировка (Вариант 1)

Евгений любил придумывать простые игры с подбрасыванием монетки. Недавно он придумал такую игру для двух человек.

Одна партия игры заключается в том, что надо кидать честную монетку до тех пор, пока не выпадет комбинация “ОХО”, где “О” — это “Орёл”, а “Х” — это либо “Орёл”, либо “Решка”. Если за 9 бросков такая комбинация выпала, то Евгений отдаёт второму игроку 1 конфету. Если же за 9 бросков такая комбинация не выпала, то второй игрок отдаёт Евгению 4 конфеты.

Помогите Евгению оценить ожидаемый выигрыш конфет после 1000 сыгранных партий игры?

В качестве ответа введите число с точностью до одного знака после запятой.

Формулировка (Вариант 2)

Евгений любил придумывать простые игры с подбрасыванием монетки. Недавно он придумал такую игру для двух человек.

Одна партия игры заключается в том, что надо кидать честную монетку до тех пор, пока не выпадет комбинация “ОХО”, где “О” — это “Орёл”, а “Х” — это либо “Орёл”, либо “Решка”. Если за 11 бросков такая комбинация выпала, то Евгений отдаёт второму игроку 1 конфету. Если же за 11 бросков такая комбинация не выпала, то второй игрок отдаёт Евгению 7 конфет.

Помогите Евгению оценить ожидаемый выигрыш конфет после 1000 сыгранных партий игры?

В качестве ответа введите число с точностью до одного знака после запятой.

Задания на анализ данных

Формулировки и решения

Файл “all_sales.csv” содержит таблицу данных с информацией о продажах продуктовых и бытовых товаров за два года. Строка данных содержит информацию о типе купленного товара (столбец “product_name”), идентификатор покупателя (столбец “client_id”) и дату покупки (столбец “date”, даты в формате Год-Месяц-День).

1. Укажите название типа товаров, который был куплен наибольшее число раз.

Ответ:

2. Укажите идентификатор покупателя, который совершил наибольшее количество покупок.

Ответ:

3. Укажите количество покупателей, совершивших всего две покупки.

Ответ:

4. Укажите количество покупателей, совершивших всего четыре покупки.

Ответ:

5. Укажите среднее число покупок за месяц.

Ответ:

6. Укажите день недели, в который за 2023 год было меньше всего куплено товаров.

Ответ:

7. Укажите день недели, в который за 2024 год было больше всего куплено товаров.

Ответ:

8. Укажите среднее число покупок за неделю (считаем, что первая неделя начинается в

понедельник 2 января 2023 года).

Ответ:

9. Укажите месяц, когда было совершено наибольшее число покупок. Формат ответа:

“Год-Месяц” (пример “2011-04”).

Укажите месяц, когда было совершено наименьшее число покупок. Формат ответа: “Год-Месяц” (пример “2011-04”).

Ответ:

10. Множество всех купленных некоторым клиентом товаров за 2 года назовём потребительской корзиной для этого клиента.

Укажите пару товаров, которая встречается в наибольшем числе потребительских корзин.

Ответ:

11. Множество всех купленных некоторым клиентом товаров за 2 года назовём потребительской корзиной для этого клиента.

Рассмотрим клиента с “client_id = 1011”. Укажите “client_id” клиента, у которого потребительская корзина имеет наибольшее пересечение с потребительской корзиной клиента с “client_id = 1011”.

Ответ:

12. Множество всех купленных некоторым клиентом товаров за 2 года назовём потребительской корзиной для этого клиента.

Рассмотрим клиента с “client_id = 1116”. Укажите “client_id” клиента, у которого потребительская корзина имеет наибольшее пересечение с потребительской корзиной клиента с “client_id = 1116”.

Ответ: