Максимальное количество баллов за олимпиаду — 72
Максимальный балл за задание — 12

Задание 1.

Модель ИИ обрабатывает запросы с двух сайтов. Вчера с первого сайта пришло на 5000 запросов больше, чем со второго. Сегодня же со второго сайта пришло на 17000 запросов больше, чем с первого. Количество запросов за оба дня с первого сайта и количество запросов за оба дня со второго сайта отличаются в два раза. Сколько всего запросов (суммарно с двух сайтов) обработала модель за два дня?

Ответ. 36000.

Критерии оценивания. Верный ответ — 12 баллов; неверный или отсутствующий ответ — 0 баллов.

Решение.

Заметим, что число запросов с сайтов отличается на $17000-5000=12000$. С другой стороны, раз эти числа отличаются в два раза, то меньшее равно 12000, а большее 24000, поэтому общее число запросов равно 36000.

Задание 2.

Разные модели отечественных самолётов имеют разное количество двигателей:

Сухой Суперджет — 2 двигателя;

МС-21 — 2 двигателя;

ТУ-154 — 3 двигателя;

ИЛ-76 — 4 двигателя.

Модель ИИ обучалась по фотографиям самолётов, которые уже введены в коммерческую эксплуатацию, поэтому она определяет по изображению тип воздушного судна (ВС) по числу двигателей следующим образом:

2 двигателя — Сухой Суперджет;

3 двигателя — ТУ-154;

4 двигателя — ИЛ-76.

Модели было предложено определить тип ВС на 100 фотографиях, на которых были изображены самолёты этих трёх моделей, а также МС-21 (который ещё в эксплуатацию не введён). На каждой фотографии был ровно один самолёт. Модель выдала 28 раз ответ «ТУ-154», 17 раз ответ «ИЛ-76», а также ровно 4 раза её ответ был неверным. На скольких из этих ста фотографий был запечатлён Сухой Суперджет?

На всех фотографиях были видны все двигатели запечатлённых на них самолётов.

Ответ. 51.

Критерии оценивания. Верный ответ — 12 баллов; неверный или отсутствующий ответ — 0 баллов.

Решение.

Модель ошибочно определяла только модель МС-21, поэтому таких изображений было 4. Итого Суперджетов $100-28-17-4=51$.

Задание 3.

В файле записаны числа 1, 2, 3, …, 999, 1000 (именно в таком порядке). Назовём число разнообразным, если в его записи есть две цифры разной чётности или в записи следующего за ним числа есть две цифры разной чётности. Например, числа 15, 18 и 1000 — разнообразные, а число 19 — нет. Сколько разнообразных чисел записано в файле?

Ответ. 984.

Критерии оценивания. Верный ответ — 12 баллов; неверный или отсутствующий ответ — 0 баллов.

Решение.

Посчитаем сначала числа, которые не являются разнообразными. У такого числа в его записи нет двух цифр разной чётности и в записи следующего за ним числа тоже нет двух цифр разной чётности.

Среди однозначных чисел подходят $1, 2, 3, 4, 5, 6, 7, 8$, их 8. Число 9 не подходит, так как следующее за ним число 10 уже содержит цифры разной чётности. Для многозначного числа: если и n, и $n+1$ не содержат двух цифр разной чётности, то при добавлении 1 должны измениться все цифры числа n. Поэтому все цифры, кроме первой, должны быть 9, а первая цифра должна быть нечётной и не равняться 9. Тогда подходят двузначные числа $19, 39, 59, 79$ и

трёхзначные числа $199, 399, 599, 799$. Число 1000 не подходит. Итого неразнообразных чисел

$$8 + 4 + 4 = 16.$$

Следовательно, разнообразных чисел

$$1000 - 16 = 984.$$

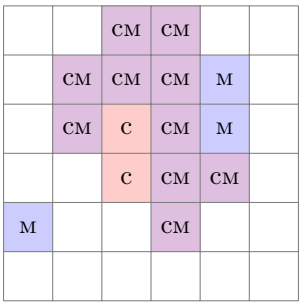
Ответ: 984.

Задание 4.

Спутниковый снимок местности разбили на квадратные клетки. Специалист отметил клетки, в которых есть возвышенности. Модель искусственного интеллекта независимо отметила клетки, которые она посчитала содержащими возвышенности. Разметку специалиста будем считать правильной. Модель может ошибаться.

Цвет и буквы внутри клетки показывают, кем она была отмечена.

Обозначение	Значение
С	Клетку отметил только специалист
М	Клетку отметила только модель
СМ	Клетку отметили и специалист, и модель



а) Сколько клеток с возвышенностями модель обнаружила правильно?

б) Сколько клеток с возвышенностями модель не отметила?

в) Сколько клеток модель ошибочно отметила как содержащие возвышенности?

г) Геологи проверят каждую клетку, которую отметили хотя бы один — специалист или модель. Сколько клеток им потребуется проверить?

д) Показатель пересечения над объединением (Intersection over Union, IoU), или индекс Жаккара, вычисляют так: число клеток, отмеченных и специалистом, и моделью, делят на число клеток, отмеченных хотя бы одним из них. Найдите IoU для данной карты. Ответ выразите в виде несократимой дроби.

Ответ. По порядку вопросов: 10; 2; 3; 15; $\frac{2}{3}$.

Критерии оценивания. За правильные ответы по порядку вопросов: 2; 2; 2; 3; 3 балла соответственно. Баллы за пункты суммируются, максимум — 12. За неверный или отсутствующий ответ на отдельно оцениваемый пункт — 0 баллов, если частичное оценивание не указано отдельно.

Решение.

На карте есть три непересекающиеся группы закрашенных клеток: фиолетовые клетки отметили оба, красные — только специалист, синие — только модель. Поэтому каждую закрашенную клетку можно отнести ровно к одной из этих трёх групп.

1. Правильно найденные клетки отмечены одновременно специалистом и моделью, то есть закрашены фиолетовым цветом. Таких клеток на карте 10.

   Ответ: 10.

2. Пропущенную клетку специалист отметил, а модель не отметила. На карте такие клетки закрашены красным цветом. Красных клеток 2.

   Ответ: 2.

3. Если клетку отметила модель, но не отметил специалист, то в действительности эта клетка не содержит возвышенности, а модель ошиблась. Такие клетки закрашены синим цветом. Их на карте 3.

   Ответ: 3.

4. Геологи проверят все закрашенные клетки: фиолетовые, красные и синие. Эти группы не пересекаются, поэтому количество клеток можно сложить:

$$10 + 2 + 3 = 15.$$

   Ответ: 15.

5. В числителе показателя IoU стоит число клеток, отмеченных и специалистом, и моделью. Это 10 фиолетовых клеток. В знаменателе стоит число клеток, отмеченных хотя бы одним из них. Как было найдено в предыдущем пункте, таких клеток 15. Поэтому

$$\text{IoU} = \frac{10}{15} = \frac{2}{3}.$$

   Ответ: $\frac{2}{3}$.

Задание 5.

В архиве хранятся пять пронумерованных сообщений. Модель получает новый запрос и отдельно сравнивает его с каждым сообщением из архива. Сам запрос в архив не добавляется.

Для каждого сообщения модель начисляет баллы по следующим правилам:

прописные и строчные буквы считаются одинаковыми, знаки препинания не учитываются;

разные формы слова считаются разными словами: например, «большой» и «большом» — разные слова;

за каждое слово, которое встречается и в запросе, и в сообщении, начисляется 1 балл.

Затем модель выбирает сообщение с наибольшим количеством баллов. Если таких сообщений несколько, выбирается сообщение с наименьшим номером.

Запрос: «кот играет на большом диване».

Сколько баллов получит каждое из пяти сообщений?

	Баллы
1. «кот спит на синем диване»	
2. «собака играет на большом диване»	
3. «большой кот сидит у окна»	
4. «кот играет с мячом»	
5. «собака спит у окна»	

Какое сообщение выберет модель?

В запрос разрешается добавить ровно одно слово, которое уже встречается в архиве. Известно, что добавляемое слово встречается ровно в одном из пяти сообщений. Какое слово нужно добавить, чтобы модель выбрала сообщение 1?

Из исходного запроса разрешается только удалять слова. После удаления слов модель снова сравнивает получившийся запрос со всеми сообщениями архива по тем же правилам. Какое наименьшее число слов нужно удалить, чтобы модель выбрала сообщение 4?

Ответ. Баллы пяти сообщений: $3, 4, 1, 2, 0$. Выбрано сообщение № 2; добавить нужно слово «синем»; минимальное число удалений — 3.

Критерии оценивания. По 0,8 балла за каждую из пяти верных оценок сообщений (всего 4); 2 балла за номер сообщения; 3 балла за слово «синем»; 3 балла за минимальное число удалений. Баллы за пункты суммируются, максимум — 12. За неверный или отсутствующий ответ на отдельно оцениваемый пункт — 0 баллов, если частичное оценивание не указано отдельно.

Решение.

1. Сравним слова запроса со словами каждого сообщения:

№	Общие слова	Баллы
1	кот, на, диване	3
2	играет, на, большом, диване	4
3	кот	1
4	кот, играет	2
5	—	0

Ответ: 3, 4, 1, 2, 0 баллов соответственно.

2. Наибольшее число баллов, равное 4, набрало сообщение № 2.

Ответ: сообщение № 2.

3. Добавим слово «синем». Оно встречается только в сообщении № 1, поэтому его результат увеличится с 3 до 4 баллов. Сообщение № 2 также имеет 4 балла, но при равенстве модель выбирает сообщение с меньшим номером.

Слово «спит» использовать нельзя: оно встречается в сообщениях № 1 и № 5 и не удовлетворяет ограничению из условия.

Ответ: «синем».

4. Удалим слова «на», «большом» и «диване». Останется запрос «кот играет», а результаты сообщений будут равны 1, 1, 1, 2 и 0 соответственно. Поэтому модель выберет сообщение № 4.

Докажем, что двух удалений недостаточно. Изначально сообщение № 2 имеет 4 балла, а сообщение № 4 — 2 балла. После двух удалений результат сообщения № 2 может уменьшиться не более чем до 2, а результат сообщения № 4 не может стать больше 2. При равенстве модель выбрала бы сообщение № 2, поскольку его номер меньше.

Ответ: 3 слова: «на», «большом» и «диване».

Задание 6.

В этой задаче модель искусственного интеллекта сравнивает слова с помощью наборов из четырёх чисел. Каждое число равно 0 или 1. Такой набор будем называть вектором смыслов. При составлении словаря каждому слову можно назначить любой такой вектор, но разным словам необходимо назначить разные векторы.

Чтобы определить, близки ли два вектора, числа на одинаковых местах перемножают, затем складывают четыре результата. Если полученная сумма больше нуля, векторы называют близкими. Если сумма равна нулю, векторы не являются близкими. Два слова называются близкими, если близки назначенные им векторы.

Например, векторы (1,0,1,0) и (0,0,1,1) близки: сумма равна 1. Векторы (1,1,0,0) и (0,0,1,1) не близки: сумма равна 0.

а) Рассмотрим все векторы из четырёх чисел, каждое из которых равно 0 или 1. Сколько из них близки к вектору $(1, 0, 1, 0)$?

б) В каких наборах любые два вектора близки?

в) Какое наибольшее число слов можно занести в словарь, назначив им векторы так, чтобы любые два слова в нём были близки?

Варианты ответа:

1. (1,0,0,0), (1,1,0,0), (1,0,1,0)

2. (1,0,0,0), (0,1,0,0), (0,0,1,0)

3. (1,1,0,0), (0,1,1,0), (1,0,1,0)

4. (1,0,0,1), (0,1,1,0), (1,1,1,1)

Ответ. 12; варианты 1 и 3; 8.

Критерии оценивания. По 4 балла за верные ответы на вопросы 1 и 3. В пункте 2: точный выбор — 4 балла; ровно одна ошибка — 2 балла; иначе — 0. Ошибка — один пропущенный верный вариант или один выбранный неверный вариант. Баллы за пункты суммируются, максимум — 12. За неверный или отсутствующий ответ на отдельно оцениваемый пункт — 0 баллов, если частичное оценивание не указано отдельно.

Решение.

1. Вектор не близок к $(1, 0, 1, 0)$, только если на первом и третьем местах у него стоят нули. Таких векторов $2 \cdot 2 = 4$, поскольку второе и четвёртое числа выбираются произвольно. Всего векторов $2^4 = 16$, поэтому близких векторов $16 - 4 = 12$.

   Ответ: 12.

2. В наборе 1 у любых двух векторов на первом месте стоит 1. В наборе 3 первая пара имеет общую единицу на втором месте, первая и третья — на первом, вторая и третья — на третьем. Эти наборы подходят. В наборе 2 первые два вектора не близки, а в наборе 4 не близки первые два вектора.

   Ответ: 1, 3.

3. Разобьём 16 векторов на 8 пар: в каждой паре один вектор получается из другого заменой всех нулей на единицы, а всех единиц на нули. Два вектора одной пары не имеют общей позиции с единицами, поэтому одновременно взять их нельзя. Значит, слов не больше 8.

   Восемь слов получить можно: возьмём все векторы, у которых на четвёртом месте стоит 1. Любые два из них близки.

   Ответ: 8.

Сириус.Курсы — для тех,
кто хочет знать больше!

