Урок 31. Вероятность и статистика в искусственном интеллекте

9 класс · Урок 31 из 34 · 45 минут

Вся программа 9 класса

Статистика в машинном обучении. На уроке для 9 класса сравниваем среднее и медиану, оцениваем вероятность и разбираем ошибки при редких событиях.

Содержание урока

- 01. Цель урока
- 02. Ход урока
- 03. Этап 1. Вступление: Почти всё верно
- 04. Этап 2. Как описать наблюдения
- 05. Этап 3. Что означает число 90%
- 06. Этап 4. Редакция ночных наблюдений
- 07. Этап 5. Проверяем вывод

01. Цель урока

Научиться описывать наблюдения числами и выбирать показатели, которые помогают оценить модель при редких событиях.

02. Ход урока

Этап	Время
1. Вступление: Почти всё верно	6 мин
2. Как описать наблюдения	11 мин
3. Что означает число 90%	8 мин
4. Редакция ночных наблюдений	15 мин
5. Проверяем вывод	5 мин
Всего	45 мин

03. Этап 1. Вступление: Почти всё верно 6 мин

ГОВОРИМ УЧЕНИКАМ

Камера школьного астроклуба сняла за ночь сто фрагментов. В десяти есть метеор, в остальных — облака, самолёты или пустое небо. Программа отмечает каждый фрагмент как «без метеора» и получает девяносто правильных ответов. Почему астроклубу от такой программы мало пользы?

Сегодня посмотрим, как статистика помогает увидеть то, что скрывается за одним числом, и от чего зависит оценка вероятности.

Для учителя

Попросите сначала письменно оценить программу одним словом. После разбора вернитесь к первым ответам и спросите, что заставило их изменить.

04. Этап 2. Как описать наблюдения 11 мин

ГОВОРИМ УЧЕНИКАМ

Сначала определите, что именно вы наблюдали. Например, пять человек проверили короткие записи и потратили 8, 9, 9, 10 и 34 секунды. Среднее время — 14 секунд, потому что сумма равна 70. Медиана — 9 секунд, середина упорядоченного списка. Если один трудный фрагмент проверяли заметно дольше остальных, среднее изменится сильнее медианы. Чтобы прикинуть, сколько займёт серия таких проверок, пригодится среднее. Медиана покажет, сколько длилась типичная проверка. Размах — разница между самым большим и самым маленьким значениями. Он покажет, насколько далеко друг от друга крайние результаты. На точечной диаграмме можно увидеть, где собралась основная часть замеров.

Статистика описывает уже полученные данные. Вероятность помогает говорить о событии, исход которого мы ещё не знаем. В этих ста фрагментах доля записей с метеором равна 10 из 100. Её можно взять как начальную оценку вероятности для случайного фрагмента из другой похожей ночи. Если отбор фрагментов изменится или начнётся метеорный поток, прежняя оценка может перестать подходить. Если же выбираем случайно именно из имеющейся сотни и все записи равновероятны, вероятность выбрать фрагмент с метеором ровно 10%.

Представьте, что клуб сохранил только записи, на которых кто-то уже заметил яркую вспышку. В такой подборке метеоров может оказаться гораздо больше, чем во всех записях камеры. Подборка пригодится, чтобы изучать вспышки, но её долю нельзя переносить на весь ночной архив. Перед оценкой выясните, как отбирали фрагменты.



Один долгий замер сдвинул среднее вправо.

Для учителя

Покажите на диаграмме два замера с одинаковым временем. Прикройте крайний замер и сравните, как изменились среднее и медиана. Затем разберите пример с ложным сигналом.

05. Этап 3. Что означает число 90% 8 мин

ГОВОРИМ УЧЕНИКАМ

Доля правильных ответов получается, когда складываем все верные отметки и делим на число записей. Из-за этого пустые кадры могут скрыть провал на редких событиях. В начале урока программа получила 90 правильных ответов, хотя пропустила все десять метеоров. Чтобы найти метеоры, отдельно посчитаем, сколько настоящих метеоров нашли и сколько пропустили. Доля найденных среди всех настоящих событий называется полнотой. А если времени на просмотр мало, важно, сколько сигналов вообще придётся открыть. Доля настоящих метеоров среди поднятых сигналов называется точностью срабатываний. У этих двух долей разные знаменатели.

Проследим за одним ложным сигналом. Он добавит работу дежурному и уменьшит долю подтверждённых сигналов. При этом число найденных метеоров не изменится, значит полнота останется прежней. С пропуском иначе: настоящий метеор остался в архиве незамеченным, поэтому полнота упадёт. Разные ошибки мешают людям по-разному.

Вопрос	Что делим	Для чего полезно
Как часто ответы верны?	Верные ответы на все проверенные записи	Общий итог проверки
Сколько метеоров удалось найти?	Найденные метеоры на все записи с метеорами	Понять, сколько редких событий пропущено

Вопрос	Что делим	Для чего полезно
Как часто сигнал подтверждается?	Найденные метеоры на все сигналы модели	Оценить, как часто просмотр сигнала будет полезен

06. Этап 4. Редакция ночных наблюдений 15 мин

ГОВОРИМ УЧЕНИКАМ

Клуб может настроить, когда программа подаёт сигнал. Допустим, модель даёт каждой записи балл от 0 до 1. Порог — это число, начиная с которого программа подаёт сигнал. Если снизить порог для тех же баллов, она сможет отметить больше записей. Среди них могут оказаться и дополнительные метеоры, и лишние сигналы. Чтобы выбрать настройку, сначала решите, какую ошибку вы готовы принять.

Задание. Выберите настройку под свою ночь

Испытали две настройки на одних и тех же ста фрагментах. Сначала обведите в таблице те числа, которые важны для вашей команды.

Настройка	Метеор найден	Метеор пропущен	Ложный сигнал	Пустой фрагмент верно отброшен
А	8	2	12	78
Б	5	5	2	88

1. Одна половина класса выбирает настройку для ночи, когда дежурный может открыть не больше восьми сигналов. Другая — для наблюдения, где клуб готов открыть до двадцати сигналов, чтобы пропустить как можно меньше метеоров.
2. Выберите настройку и запишите, какой ценой получаете нужный результат. Обменяйтесь решениями и объясните, почему более высокая доля верных ответов может оказаться слабым аргументом.
3. Предложите, какие новые записи надо проверить до следующего запуска. Назовите одно условие, при котором прежние числа могут измениться.

Результат: решение для своей ночи, подтверждение двумя числами и предложение для новой проверки.

Для учителя

А даёт 20 сигналов, находит 8 из 10 и имеет 86 правильных ответов из 100. Б даёт 7 сигналов, находит 5 из 10, получает 93 верных ответа. Доля настоящих метеоров среди сигналов: 8/20 и 5/7. При лимите 8 подходит Б, при лимите 20 и приоритете

полноты — А. Для новой проверки полезны записи при другом освещении или с большим числом метеоров, отдельно от настройки порога.

07. Этап 5. Проверяем вывод 5 мин

ГОВОРИМ УЧЕНИКАМ

Одна доля описывает один вопрос. Перед сравнением моделей скажите, что именно входит в числитель и знаменатель и для каких условий собирали данные.

Главное за урок

- Медиана показывает середину наблюдений, а среднее учитывает величину каждого значения.
- Частота на прошлых данных помогает оценить вероятность в сходных условиях.
- Для редких событий важно отдельно проверять пропуски и ложные сигналы.

Итоговые вопросы

1. Один из пяти замеров удвоился. Как узнать, какой показатель изменится сильнее, и почему?
2. После смены камеры обнаружились новые блики. Зачем снова проверять модель?

Для учителя

1. Если удвоился крайний большой замер, среднее вырастет, а медиана не изменится. При изменении другого числа надо сначала перестроить порядок значений. 2. Условия съёмки изменились, поэтому прежние оценки ложных сигналов и пропусков могут больше не подходить.

Источники

- NIST. Среднее, медиана и другие показатели положения данных · Проверено 05.11.2026. Примеры показывают, как крайние значения могут заметно сдвинуть среднее относительно медианы.

- Google. Доля верных ответов, полнота и точность срабатываний · Проверено 05.11.2026. Показатели рассчитываются для заданного порога и меняются вместе с ним.