Урок 29. Линейная алгебра и векторы в искусственном интеллекте

9 класс · Урок 29 из 34 · 45 минут

Вся программа 9 класса

Векторы в искусственном интеллекте. На уроке для 9 класса записываем признаки числами и проверяем, как масштаб и веса влияют на поиск похожих объектов.

Содержание урока

- 01. Цель урока
- 02. Ход урока
- 03. Этап 1. Вступление: Две похожие игры
- 04. Этап 2. Как числа описывают предмет
- 05. Этап 3. Сделайте две версии поиска
- 06. Этап 4. Чего нет в векторе
- 07. Этап 5. Проверяем понимание

01. Цель урока

Понять, как математика помогает представить предмет числами и сравнить его с другими предметами по заданным признакам.

02. Ход урока

Этап	Время
1. Вступление: Две похожие игры	5 мин
2. Как числа описывают предмет	10 мин
3. Сделайте две версии поиска	17 мин
4. Чего нет в векторе	8 мин
5. Проверяем понимание	5 мин

Этап	Время
Всего	45 мин

03. Этап 1. Вступление: Две похожие игры 5 мин

ГОВОРИМ УЧЕНИКАМ

Клуб настольных игр хочет подобрать замену игре, которую уже забрали. Есть две коробки: одна на то же число игроков, другая на то же время. Похожи они на оригинал по разным причинам. Что должен знать поиск, чтобы предложить подходящую замену?

Сегодня научимся описывать предметы числами и задавать, что именно считаем сходством. Это помогает понять, почему векторы так часто встречаются в искусственном интеллекте.

Для учителя

Соберите две версии. Спросите, как изменится выбор, если лишний человек останется без места или если у группы всего полчаса.

04. Этап 2. Как числа описывают предмет 10 мин

ГОВОРИМ УЧЕНИКАМ

Начнём с того, что важно для выбора игры. Можно записать число игроков, длительность в минутах и условную оценку количества разговоров. Получится вектор, например (4, 25, 2). Вектор признаков — это упорядоченный набор чисел. Мы знаем, какую характеристику означает каждое место. Если поменять числа местами, запись будет описывать другую игру или окажется бессмысленной.

Когда нужно искать похожие вещи, программа может сравнивать значения по координатам. Допустим, разница в длительности равна десяти минутам, а разница в числе игроков — двум. Если просто сложить эти числа, минуты будут влиять сильнее. Можно договориться, что для сравнения один шаг — это один игрок, пять минут или один балл по шкале общения. Тогда 10 минут дают две единицы различия: 10 ÷ 5 = 2. Затем каждой разнице можно дать вес, то есть множитель, который показывает её важность. Для сравнения векторов есть разные способы. В этом уроке будем складывать взвешенные различия: чем меньше сумма, тем ближе игра к запросу.

Посмотрим на две строки ниже. У игр совпадает число участников. По времени разница в два шага, по шкале общения тоже в два. Если оба признака равноценны, сумма равна четырём. А если один шаг по шкале общения считаем вдвое важнее, получаем 2 + 2 × 2 = 6. Сами игры не изменились. Изменился способ сравнения. По этой же причине каталог материалов может по-разному подбирать статью для быстрого обзора и подробного чтения.

Линейная алгебра даёт язык для таких задач. Векторы хранят координаты, признаки или параметры. В матрицу можно собрать много векторов с одинаковым порядком признаков. Тогда удобнее обрабатывать записи сразу. А система линейных уравнений нужна, когда одно и то же неизвестное должно удовлетворять нескольким условиям. Например, по общему весу двух разных наборов можно найти вес каждой детали, если измерений достаточно. В моделях та же математика помогает задавать признаки, подбирать параметры и проверять вычисления.

Запись	Первое число	Второе число	Третье число
(4, 25, 2)	4 игрока	25 минут	Мало разговоров, 2 из 5
(4, 35, 4)	4 игрока	35 минут	Много разговоров, 4 из 5

Для учителя

Прочитайте координаты и вместе с классом посчитайте разницу во времени. Затем покажите, как вес меняет сумму, и перейдите к реальному случаю.

Реальный случай. Поиск переводов по векторам

В июле 2020 года исследователи опубликовали работу о LaBSE. Эта модель превращала предложения на разных языках в числовые векторы. Если запрос и его перевод оказываются рядом, система может найти перевод среди множества фраз. Авторы проверяли такой поиск на наборе Tatoeba, в котором есть переводы предложений. Это пример того, как с помощью чисел можно сравнивать тексты, даже когда в них разные слова и алфавиты.

05. Этап 3. Сделайте две версии поиска 17 мин

ГОВОРИМ УЧЕНИКАМ

В каталоге можно оставить один и тот же набор признаков, но по-разному настроить поиск. Сначала договоритесь, что важнее для вашего посетителя. Тогда будет понятно, почему алгоритм выдал именно такую рекомендацию.

Задание. Две группы, один каталог

Вектор запроса обеих групп — (4, 25, 2). В этом каталоге игры рассчитаны ровно на указанное число участников. При сравнении одна единица разницы означает одного игрока, пять минут или один балл по шкале общения.

Игра	Игроки	Минуты	Шкала общения
Компас	4	35	2
Шифр	3	25	3
Развязка	4	30	4
Черта	5	20	2

1. В паре придумайте для каждой группы условие, при котором поиску полезно поставить на первое место разные игры. Объясните, почему условие важно.
2. Для одной группы выберите обязательное ограничение либо веса и сравните минимум два подходящих варианта. Запишите, как меняется выбор, если важное условие убрать.
3. Передайте соседям исходный запрос, ваши условия и веса без названия выбранной игры. Пусть они восстановят ваш выбор и укажут место, где правила оставляют сомнение.

Результат: две рекомендации, способ сравнения и одна проверка правил чужой парой.

Для учителя

Можно выбрать «Компас», если важны четыре места и тихая игра, а 35 минут доступны. «Развязка» подходит для четырёх человек, если важнее уложиться по времени, чем выбрать тихую игру. Если позже придёт группа из пяти человек, она задаст новый запрос. Для неё при лимите 25 минут можно предложить «Черту». Если участник не может присоединиться, число мест следует проверять как обязательное условие до сравнения. При весах (3, 1, 2) до отбора по числу мест получатся суммы 2 для «Компаса», 5 для «Развязки», 4 для «Черты» и 5 для «Шифра». Если нужны ровно четыре места, сравнивайте только «Компас» и «Развязку». Другие обоснованные веса допустимы.

06. Этап 4. Чего нет в векторе 8 мин

ГОВОРИМ УЧЕНИКАМ

Если две игры получили одинаковый вектор, выбранный способ их не различит. Например, в каталоге нет сведений о том, играют ли участники друг против друга или всей командой. Можно добавить новый признак, если он помогает людям выбирать игру. Придётся решить, как одинаково записывать его для всех игр.

Обсуждение. Что стоит измерить

Назовите ещё одно свойство, которое меняет решение вашей группы. Предложите, как записать его числами, и придумайте два случая, где это свойство действительно меняет рекомендацию. Обсудите, можно ли измерить его одинаково для двух разных игр.

Для учителя

Например, режим «совместно» можно записывать как 1, «нет» как 0, а игры с двумя режимами отмечать отдельным полем. Если ученики предложат «интересность», попросите придумать более ясный способ наблюдения или заменить признак.

07. Этап 5. Проверяем понимание 5 мин

ГОВОРИМ УЧЕНИКАМ

Когда рекомендация выглядит странно, сначала проверьте, что программа знала об объекте и как сравнила значения. Этого часто достаточно, чтобы найти причину.

Главное за урок

- Координата вектора имеет смысл, если заранее известно, какой признак она описывает.
- Единицы и веса меняют то, какие объекты окажутся похожими.
- Модель может сравнить только те различия, которые доступны в её представлении данных.

Итоговые вопросы

1. Почему (4, 25, 2) и (25, 4, 2) описывают разные вещи?
2. Если ведущему нужна игра ровно на четыре места, когда лучше отсеять неподходящие игры?
3. Почему два разных предложения могут иметь близкие векторы в системе поиска переводов?

Для учителя

1. У координат закреплён порядок. 2. До ранжирования по остальным признакам. 3. Модель училась сближать представления фраз с похожим смыслом на разных языках.

Источники

- Feng и соавторы. Language-agnostic BERT Sentence Embedding · 03.07.2020, редакция 08.03.2022, материал старше 2024 года. «bi-text retrieval accuracy over 112 languages on Tatoeba» — проверка поиска соответствующих предложений.
- Google. Приведение числовых признаков к сопоставимому масштабу · Проверено 05.11.2026. Подход к нормализации выбирают с учётом того, как распределены числовые значения.