Study Space Recommendation System

Problem

Students have different preferred environments for studying.

So we can use math to help / recommend them study spots.

They filled out a survey that asks them a bunch of questions to get a better understanding of their preffered study space.

Translate mathematically

Here we walk through a simpler problem where we ask them about 2 features so we ask their (1) preferred noise level, (2) preferred crowd level

We also have simple fake data for this problem

Study Space	Noise (1-5)	Crowd (1-5)
Library	1	2
Cafe	4	4
Lounge	3	2

We turn each study space to vectors

$$\vec{v}_{library} = \begin{bmatrix} 1 \\ \cdot \end{bmatrix}, \quad \vec{v}_{cafe} = \begin{bmatrix} 4 \\ \cdot \end{bmatrix}, \quad \vec{v}_{lounge} = \begin{bmatrix} 3 \\ \cdot \end{bmatrix}$$

$$\vec{v}_{library} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \vec{v}_{cafe} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}, \vec{v}_{lounge} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Let's say a student named Studi filled out the survey

$$\text{Studi} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

This tells us Studi wants 2 for noise level and 1 for crowd level.

So how do we recommend study spaces based on Studi's inputs and the data we have. We'll use distance formula / Euclidean norm

Library: $\sqrt{(2-1)^2+(1-2)^2} = \sqrt{2}$

Cafe: $\sqrt{(2-4)^2+(1-4)^2} = \sqrt{13}$

Lounge: $\sqrt{(2-3)^2+(1-2)^2} = \sqrt{2}$

We look at the shortest distance from Studi to a study space. Here we see that library and lounge are the shortest distances so we recommend both for Studi.

Why this works?

In the example, we want to recommend Studi the closest noise level and crowd level based on our data and their input. Distance formula is the math for that

In general

[s. 7

In general

Each study space will have n features $\underline{s} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} \in \mathbb{R}^n$

Each student will also have n inputs $\underline{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{R}^n$

We calculate distance for each student to each study space.

$$d(u,s) = \sqrt{(u_1 - s_1)^2 + \ldots + (u_n - s_n)^2} = \sqrt{\sum_{i=1}^{n} (u_i - s_i)^2}$$

Then we determine the shortest distance/s and recommend those study space/s.

In Python, we can use KNN algorithm to do that.