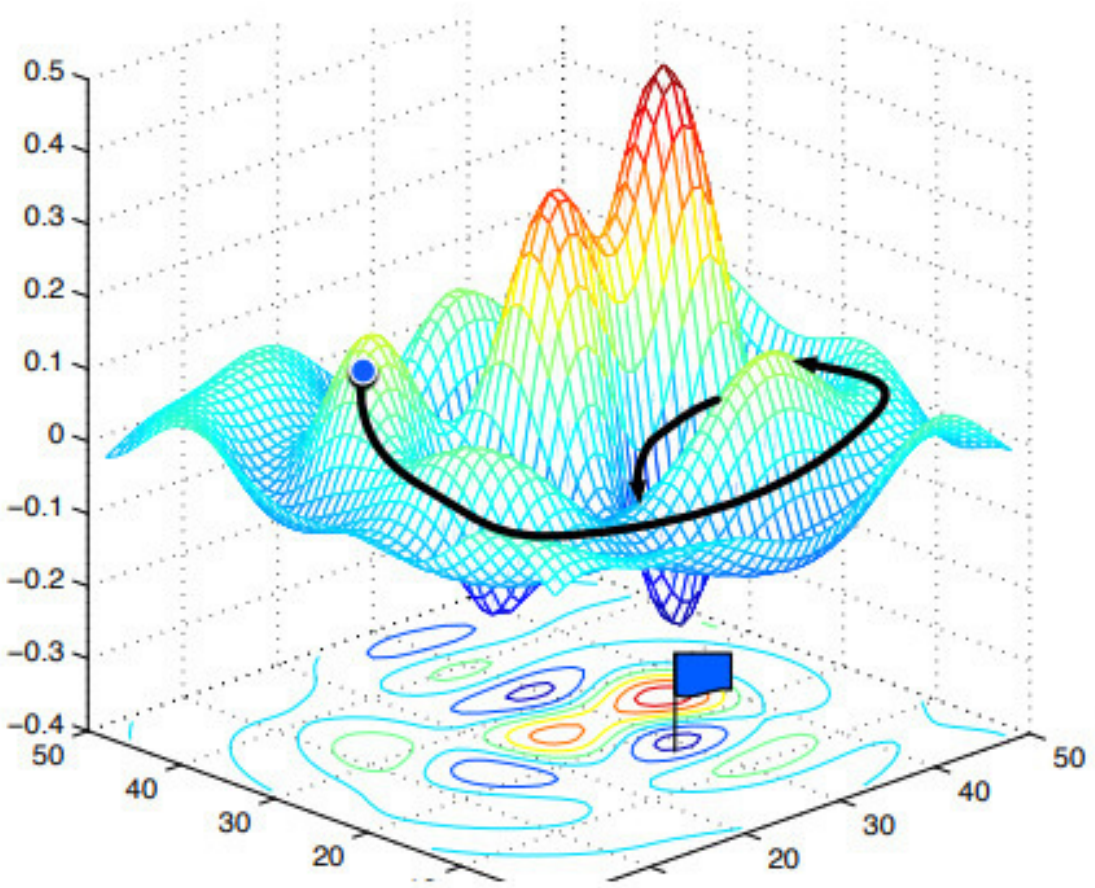
Gradient descent 101 - cookies and local minima



anusha-m

Ever gotten lost in a mall trying to find the food court?

anusha-m

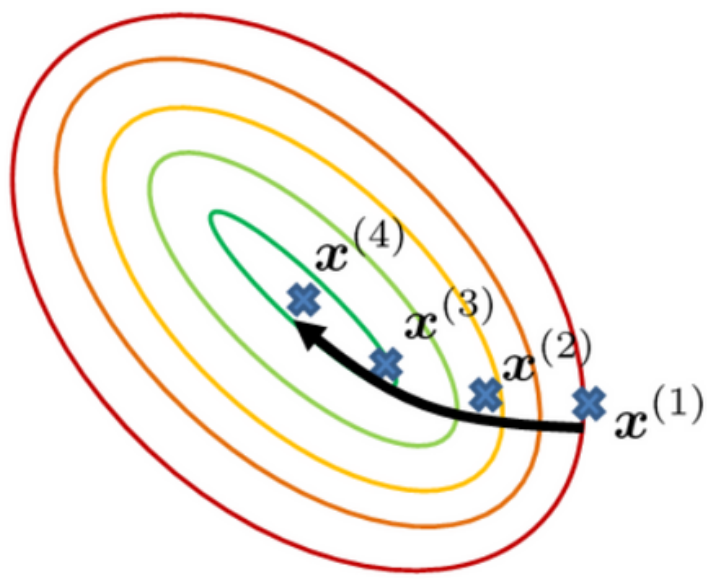
You enter the mall with a clear objective but the irresistable smell of cookies leads you in a different direction.

You follow the sweet smell, and it brings you to a bakery on the ground floor. You haven't found the food court. Turns out it's on the top floor!

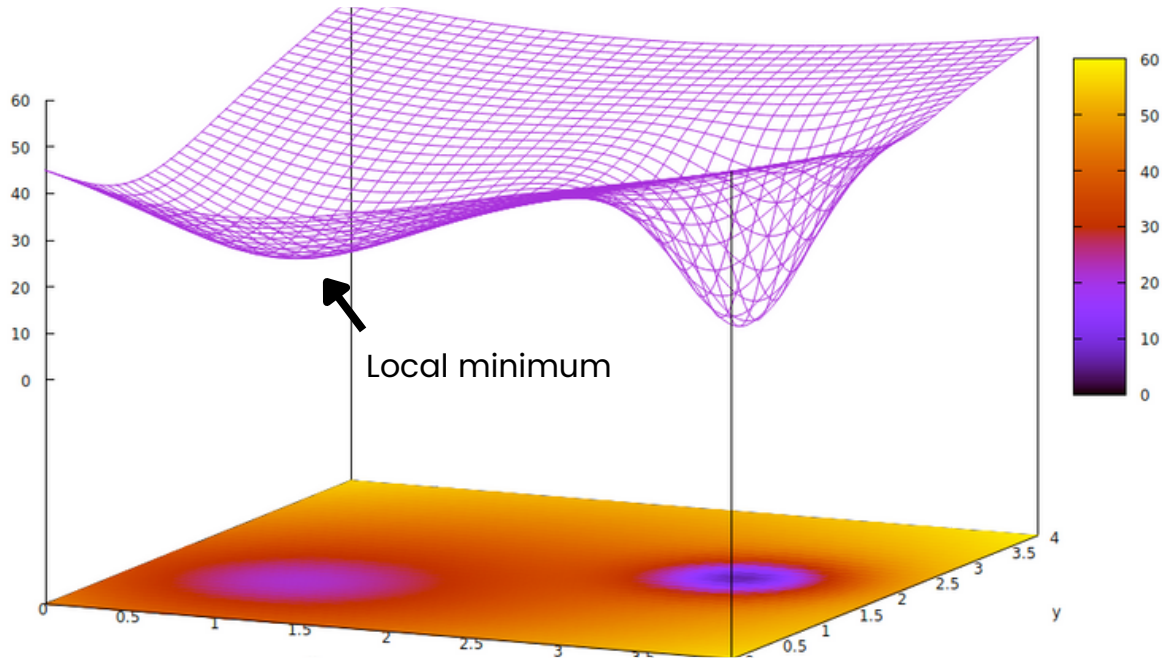
This scenario represents the problem of local minima in gradient descent, a commonly used optimization algorithm in machine learning.

anusha-m

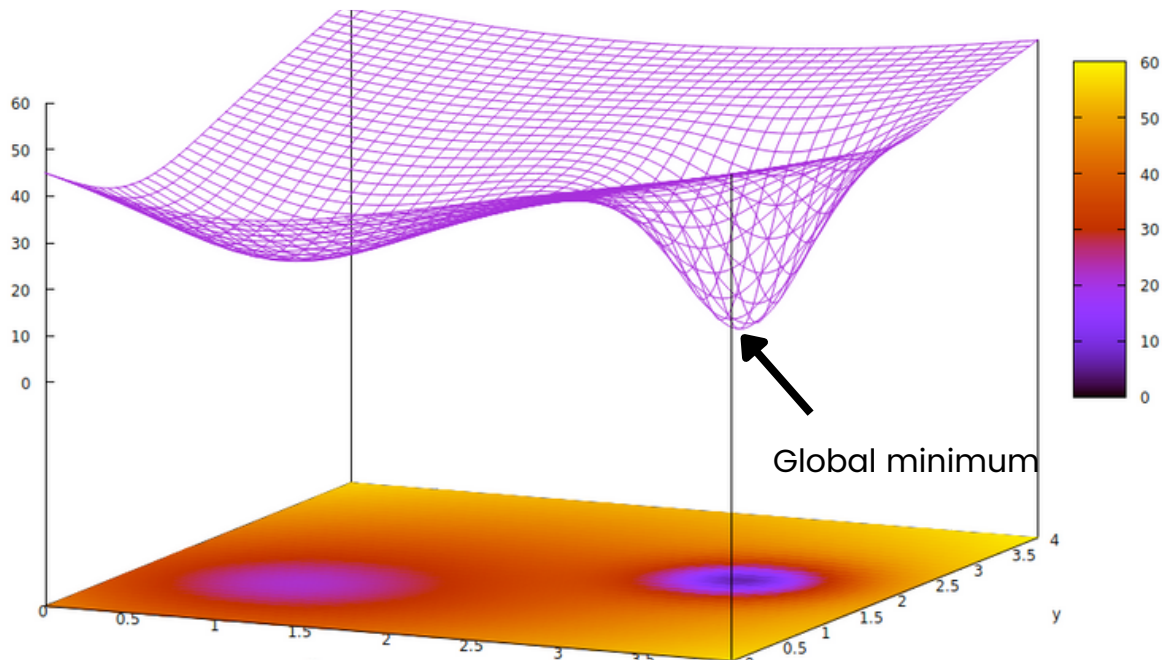
The smell of cookies represents the "pull", or the slope. Naïvely following the smell represents the greedy approach adopted by gradient descent.



The bakery you ended up in is a local minimum, which means it's the optimal point in your immediate surroundings but not the optimal point overall.



The food court on the top floor is the global minimum - the ultimate objective.

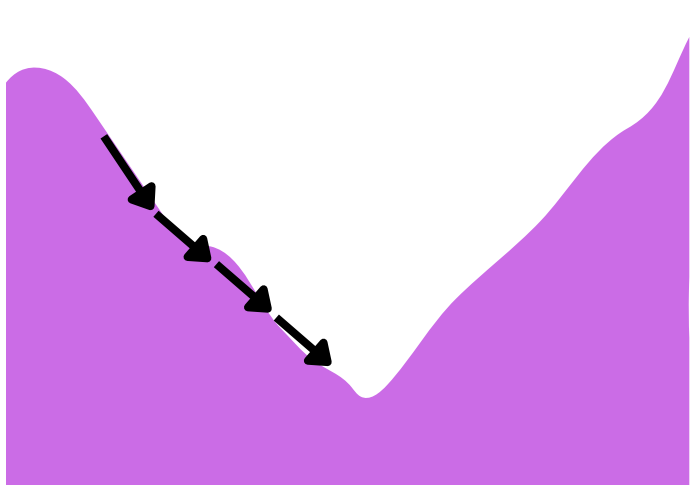


anusha-m

Applying gradient descent in machine learning involves updating the parameters of a model in the direction of steepest descent of the loss function.

anusha-m

Just like you unwaveringly followed the smell of the cookies, gradient descent follows the slope of the loss function to reach the optimal solution.



anusha-m

However, this approach can often lead to a local minimum (especially in non-convex functions) like the bakery, which is not what we were looking for at all.

To overcome the problem of local minima, techniques such as random restarts and simulated annealing are used to explore different parts of the loss function and eventually reach the global minimum.

anusha-m

So the next time you either get lost in a mall or are faced with an optimization problem, remember the importance of exploring different paths and not getting stuck in a local minimum.

anusha-m

Keep your eyes on the global minimum and don't let the smell of cookies distract you!

What's your favourite optimization algorithm and why? Let me know ↓

anusha-m