Gossip algorithm in deeplearning .. like the one in distributed system.

Remember:

If we always pick the new local optimum , we’re doing a random walk (a sort of meta-exploration). If we only pick the new local optimum if it’s better than our current one, we’re doing hill-climbing (a sort of meta-exploitation). ILS often picks something in-between the two, as discussed later.