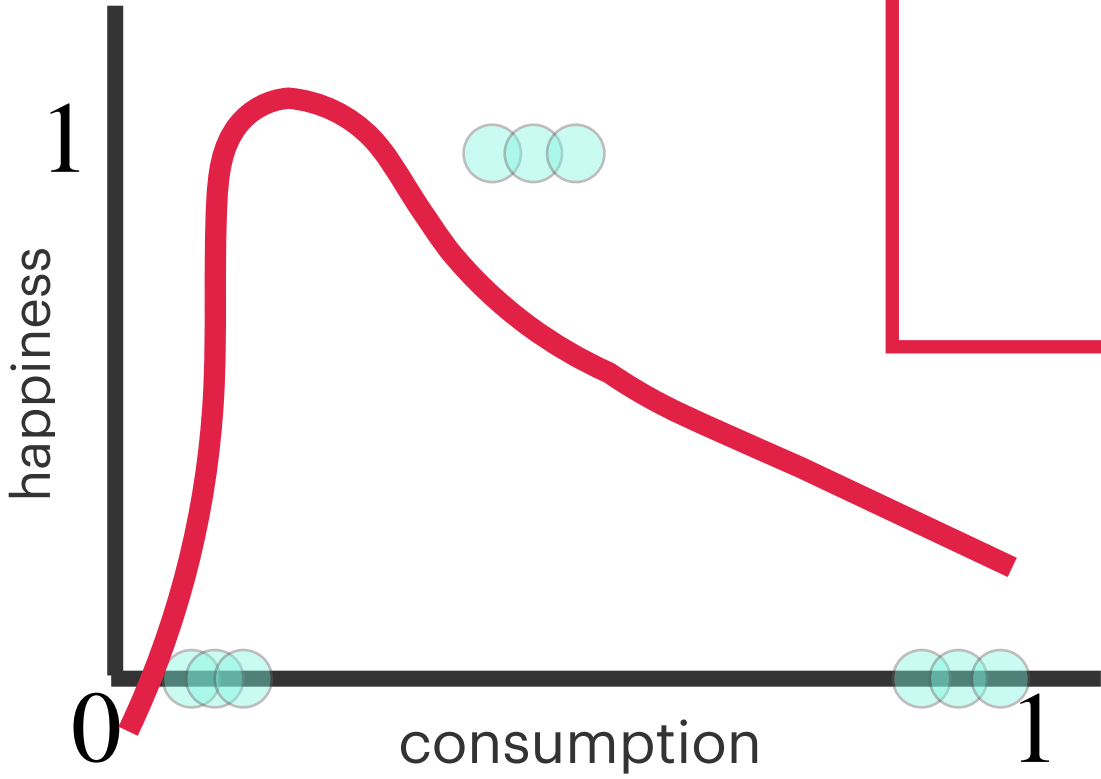
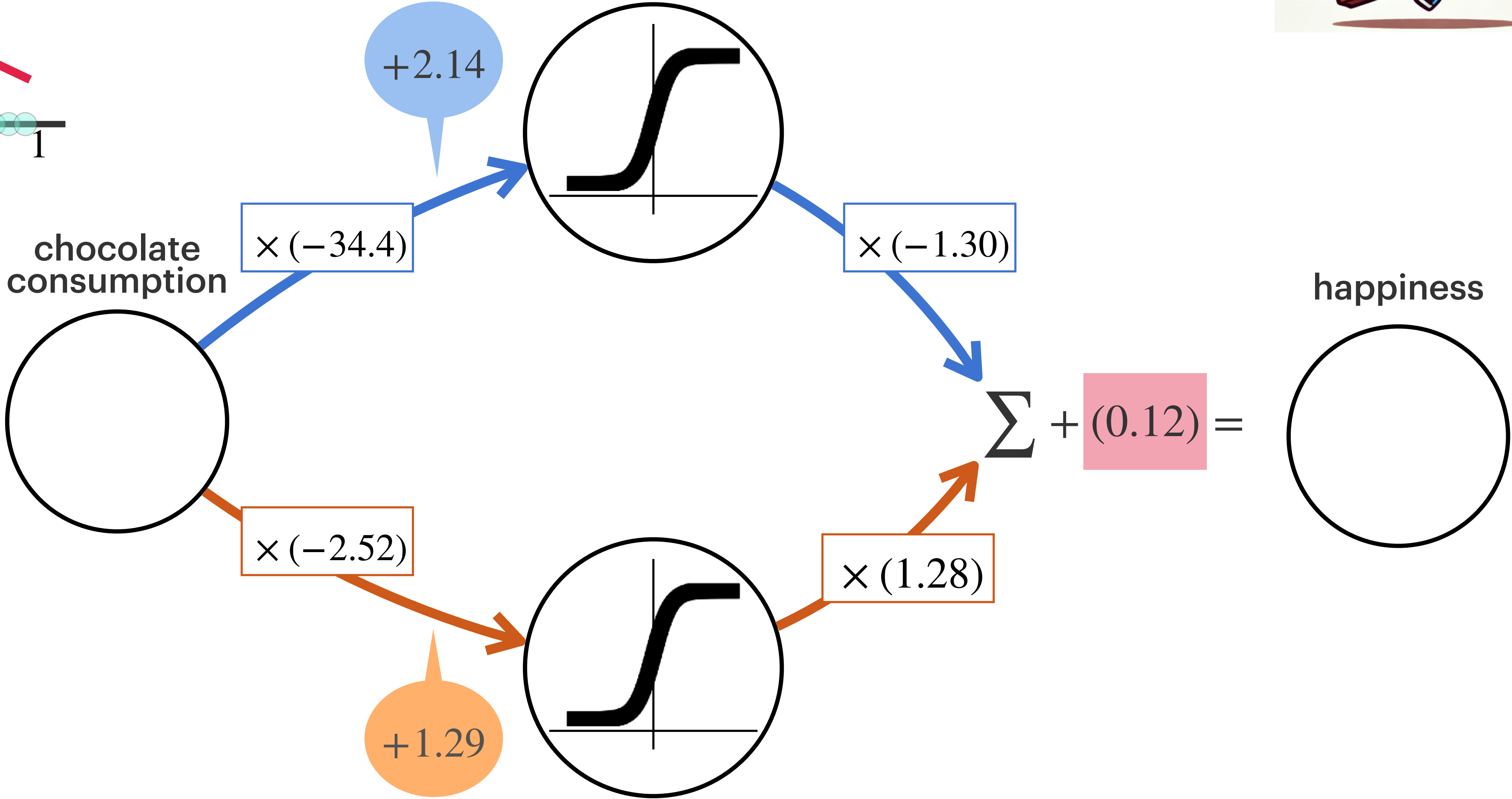


# How Do Neural Networks Work?



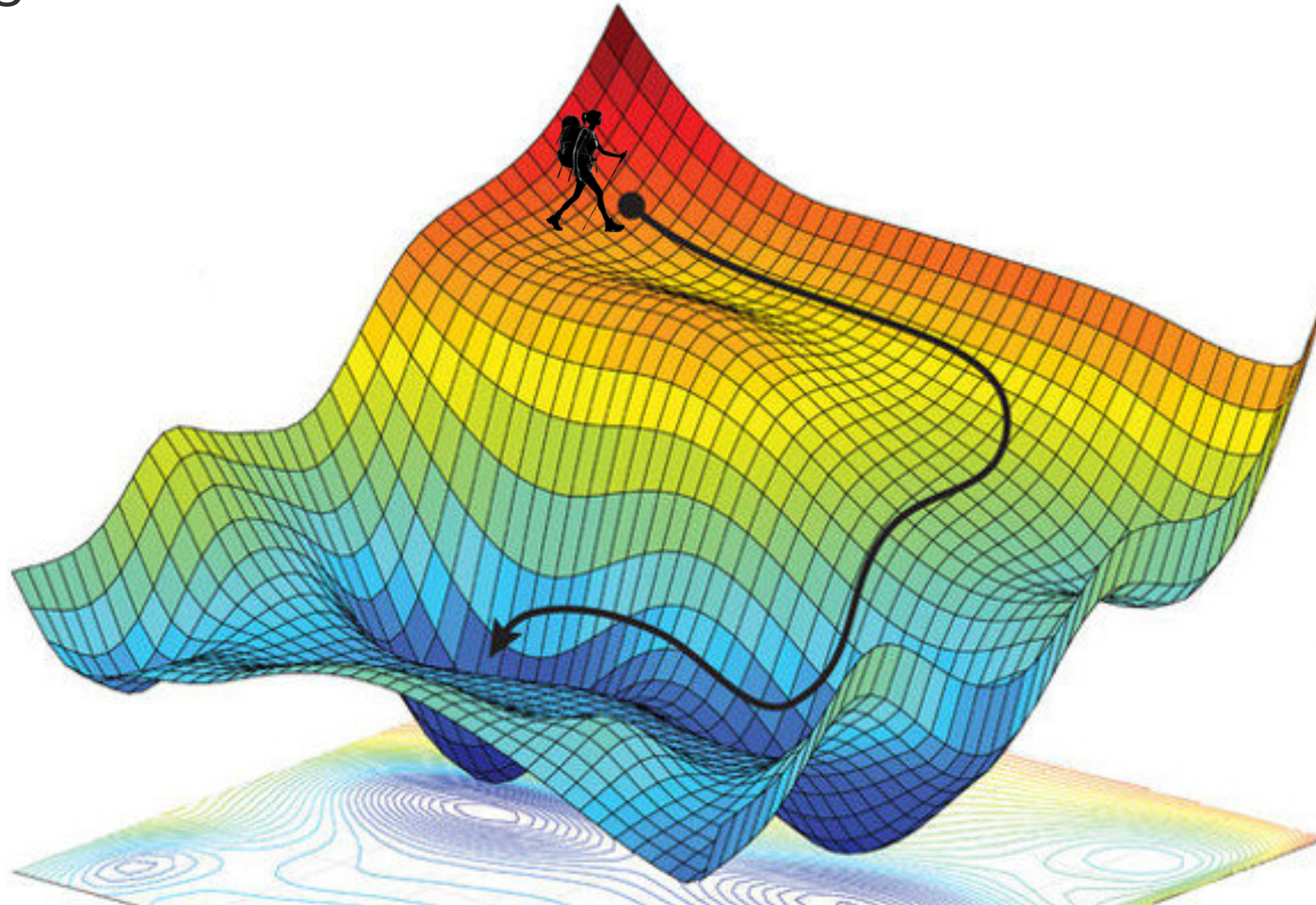
and then add 0.02 to get the final curve





# Estimation: Gradient Descent

- The goal is to minimize the loss function
- The gradient tells us which direction to move



[source: <https://arxiv.org/abs/1805.04829>]