Gemini Chat

- https://gemini.google.com/share/7e34876ac7cb

در این تمرین به طور اشتباه از چت قبلی استفاده کرده ام.

Notebook LM

- https://notebooklm.google.com/notebook/70ec6ac4-b5af-4d59-8554-178567f6ea1d

References

- https://arxiv.org/abs/1412.6980
- https://www.datacamp.com/tutorial/adam-optimizer-tutorial
- https://en.wikipedia.org/wiki/Stochastic_gradient_descent
- https://www.kaggle.com/code/tanujasreekanth/comparing-optimizers-fashion-mnist-data-set
- https://jamhuri.medium.com/understanding-the-adam-optimization-algorithm-a-deep-dive-into-the-formulas-3ac5fc5b7cd3
- https://medium.com/mini-distill/effect-of-batch-size-on-training-dynamics-21c14f7a716e
- https://www.reddit.com/r/math/comments/e72q34/the_concept_of_convergence_in_gradient_descent/
- https://sh-tsang.medium.com/brief-review-visualizing-the-loss-landscape-of-neural-nets-dd93cb261afc