



MITx 6.86x

Machine Learning with Python-From Linear Models to Deep Learning[Course](#)[Progress](#)[Dates](#)[Discussion](#)[Resources](#)[Course](#) / [Unit 3. Neural net...](#) / [Lecture 9. Feedforward Neural Networks, Back](#)[< Previous](#)

3. Training Models with 1 Hidden Layer

[Bookmark this page](#)

Exercises due Mar 29, 2023 08:59 -03 Completed

Video note: In the video below at 0:38, Prof Jaakkola misspoke and said "on the left," but the plot is "on the **right**".

Training Models with 1 Hidden Layer, Overcapacity, and Convergence Guarantees




▶ 0:00 / 0:00


▶ 1.0x

Video

 [Download video file](#)

Transcripts

 [Download SubRip \(.srt\) file](#)

 [Download Text \(.txt\) file](#)

SGD Convergence guarantees

1/1 point (graded)

Which of the following option(s) is/are true about training neural networks?

(Choose all that apply.)

☐ For multi-layer neural networks, stochastic gradient descent (SGD) is guaranteed to converge to a global optimum

☒ For multi-layer neural networks, stochastic gradient descent (SGD) is not guaranteed to converge to a global optimum

Legal

[< Previous](#)[Next >](#)[Terms of Service & Honor Code](#)[Privacy Policy](#)[Accessibility Policy](#)[Trademark Policy](#)[Sitemap](#)[Cookie Policy](#)[Do Not Sell My Personal Information](#)

Connect

[Blog](#)[Contact Us](#)[Help Center](#)[Security](#)[Media Kit](#)

© 2023 edX LLC. All rights reserved.

深圳市恒宇博科技有限公司 [粤ICP备17044299号-2](#)