

EdX and its Members use cookies and other tracking technologies for performance, analytics, and marketing purposes. By using this website, you accept this use. Learn more about these technologies in the [Privacy Policy](#).



MITx: 6.86x

Machine Learning with Python-From Linear Models to Deep Learning

[Help](#)[sandipan_dey](#)

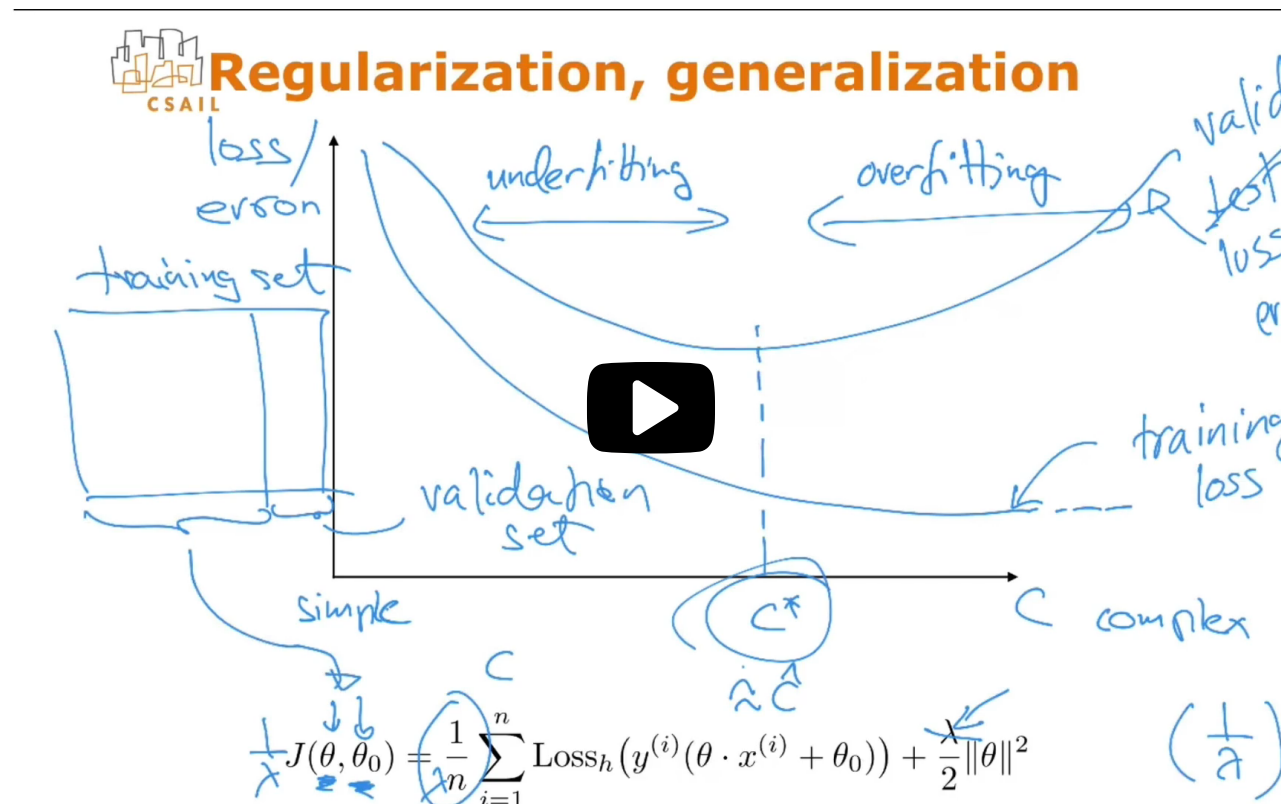
[Unit 1 Linear Classifiers and](#)
[Course](#) > [Generalizations \(2 weeks\)](#)

[Lecture 4. Linear Classification and](#)
> [Generalization](#)

> 3. Regularization and Generalization

3. Regularization and Generalization

Regularization and Generalization



and use the validation examples to evaluate an approximate test

and error.

So instead, what we are evaluating here

is not actually the test loss and test error.

We are evaluating the large validation error.

And then we find the value of c that

actually optimizes the performance

on those pretend test examples.

So we don't get c star exactly.

But we get some estimate approximate value of c star.

▶ 7:07 / 7:07

▶ Speed 1.50x



Video

[Download video file](#)

Transcripts

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

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

Generalization 1

2/2 points (graded)

If the training loss is low and the validation loss is high, the model might be:

☐ underfitting

☒ overfitting ✓

☐ fits well

If the training loss is high and the validation loss is high, the model might be:

☒ underfitting ✓

☐ overfitting

☐ fits well

Solution:

If the model is doing very well on the training set but perform purely on the validation set, it means that it learned features that are very specific for the training set and that are not general enough.

Submit

You have used 1 of 1 attempt

i Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 1 Linear Classifiers and Generalizations (2 weeks):Lecture 4. Linear Classification and Generalization
/ 3. Regularization and Generalization

Add a Post

Show all posts ▼		by recent activity ▼	
✓	<u>Significance of C</u> I don't understand what is the significance of dividing objective function by lambda? Can't we just use lambda thus no need of C?	6	▼
💬	<u>confusion. 1st question , wording , lost point</u> When training loss is low : it means , the loss got exponentially lower to minimal . Validation loss large means the model can't properly classify . It means the model is bad and...	2	▼
💬	<u>The Meaning of Complexity in the context of Linear Classifiers</u> The Professor equates a large $ \theta $ with complexity, but he does not define what he means by complexity. Generally speaking complexity refers to the capacity of the model to...	5	▼
💬	<u>SVM - why not cap the hinge loss in the objective function?</u>	11	▼
✓	<u>Can't understand pharase.</u> Hi. About 2:04 prof. says : > Notice that it is monotonically decreasing the more. I emphasize the > **training bars.** In the objective function, the lower the value will be. Wha...	5	▼
💬	<u>Error in transcript</u> We are evaluating the large validation error. At the end of transcript. It should be the following instead: We are evaluating validation loss and validation error.	3	▼

Learn About Verified Certificates