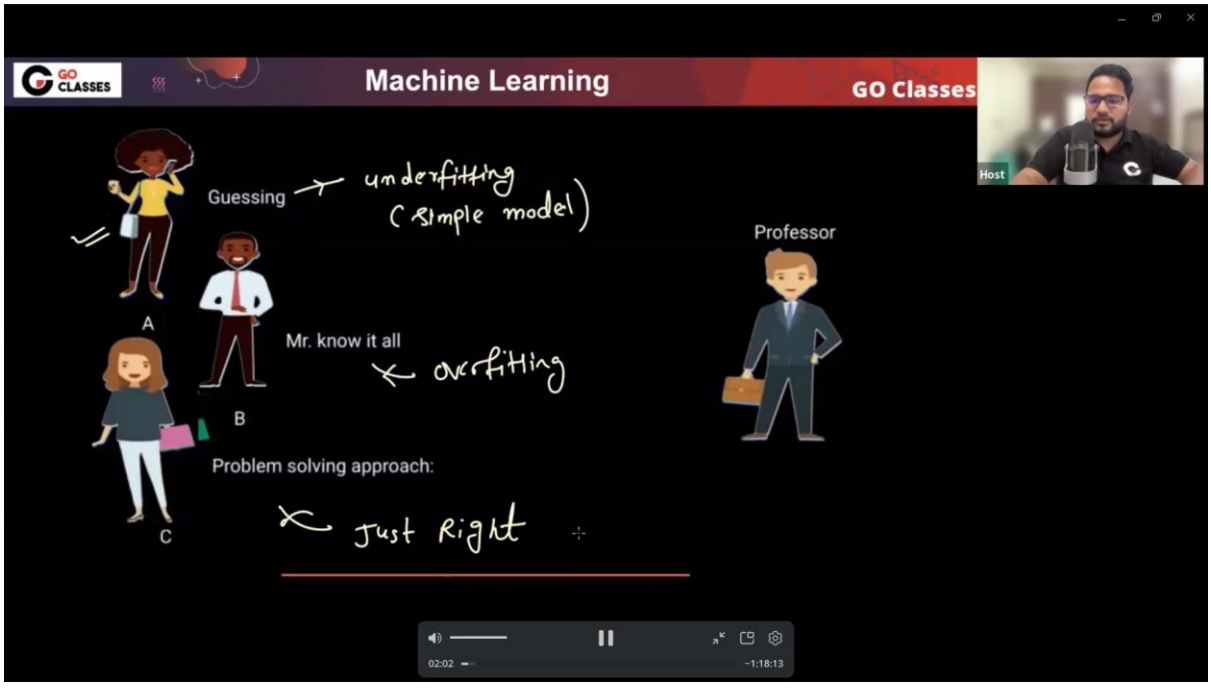
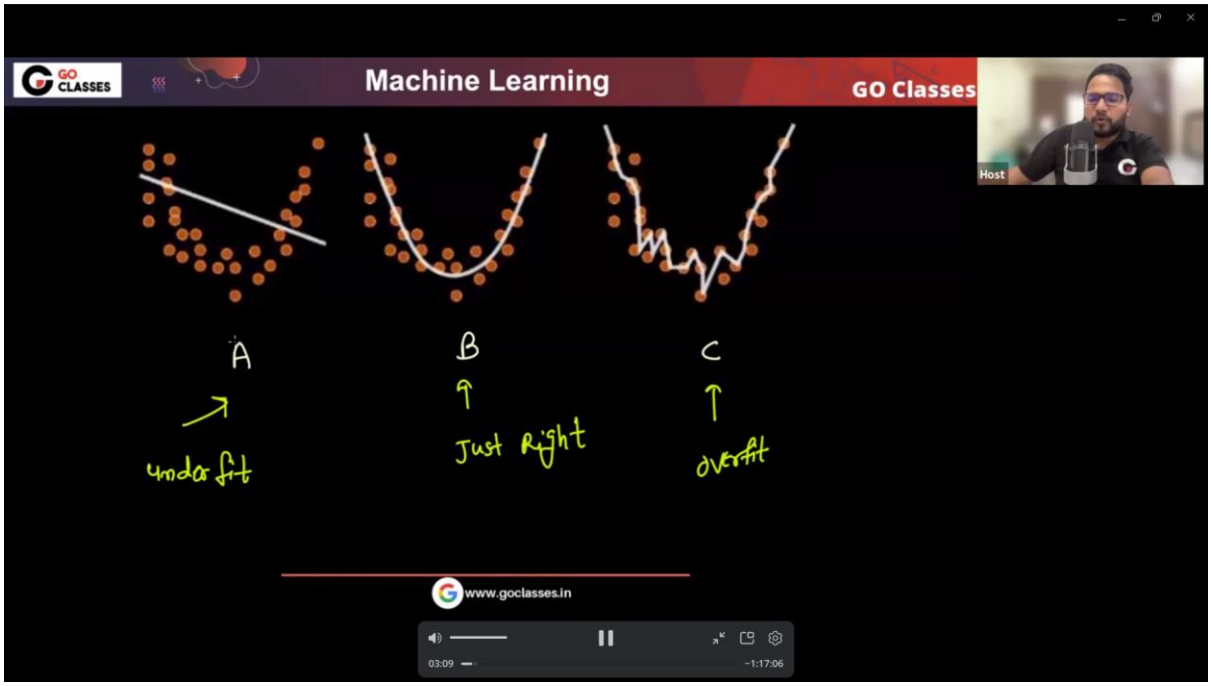
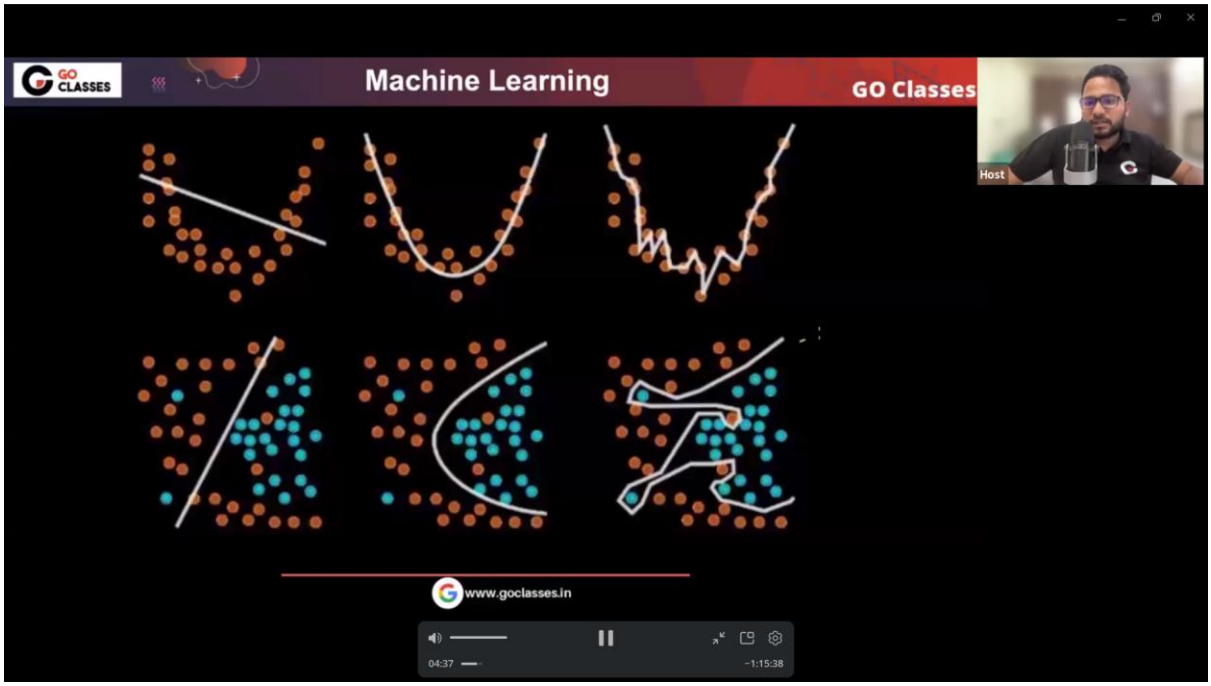
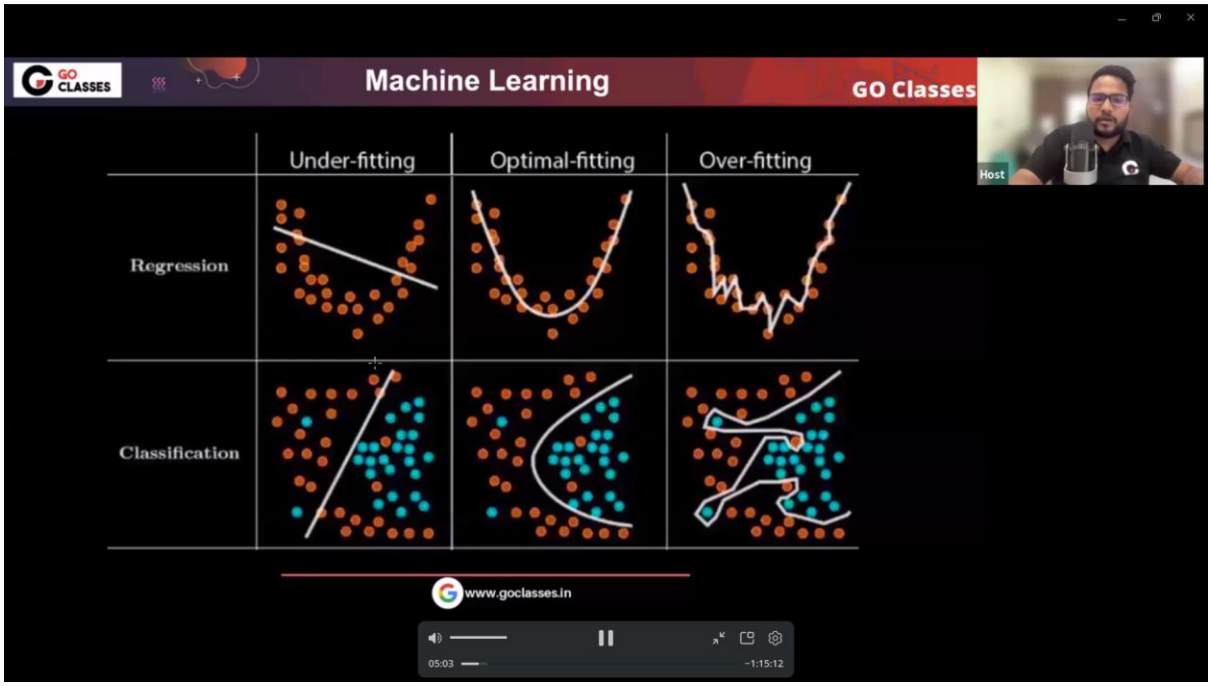
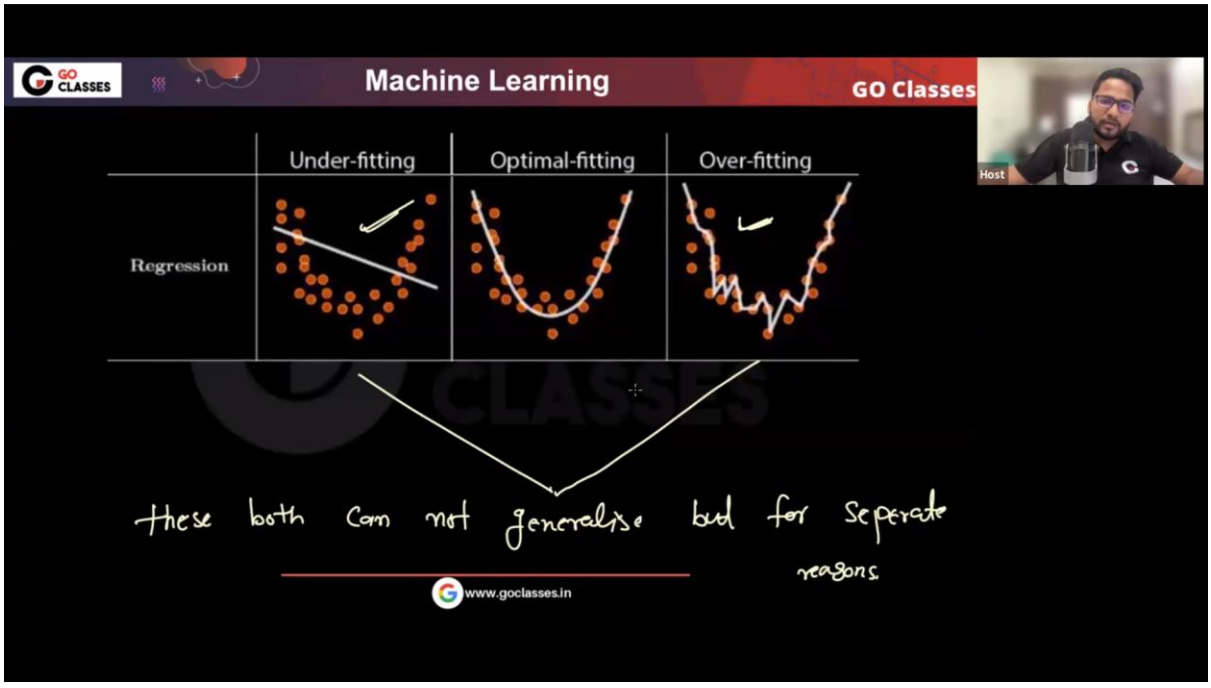

GO CLASSES
Machine Learning
GO Classes
Host
Bias-Variance Tradeoff
www.goclasses.in
00:43
-1:19:32

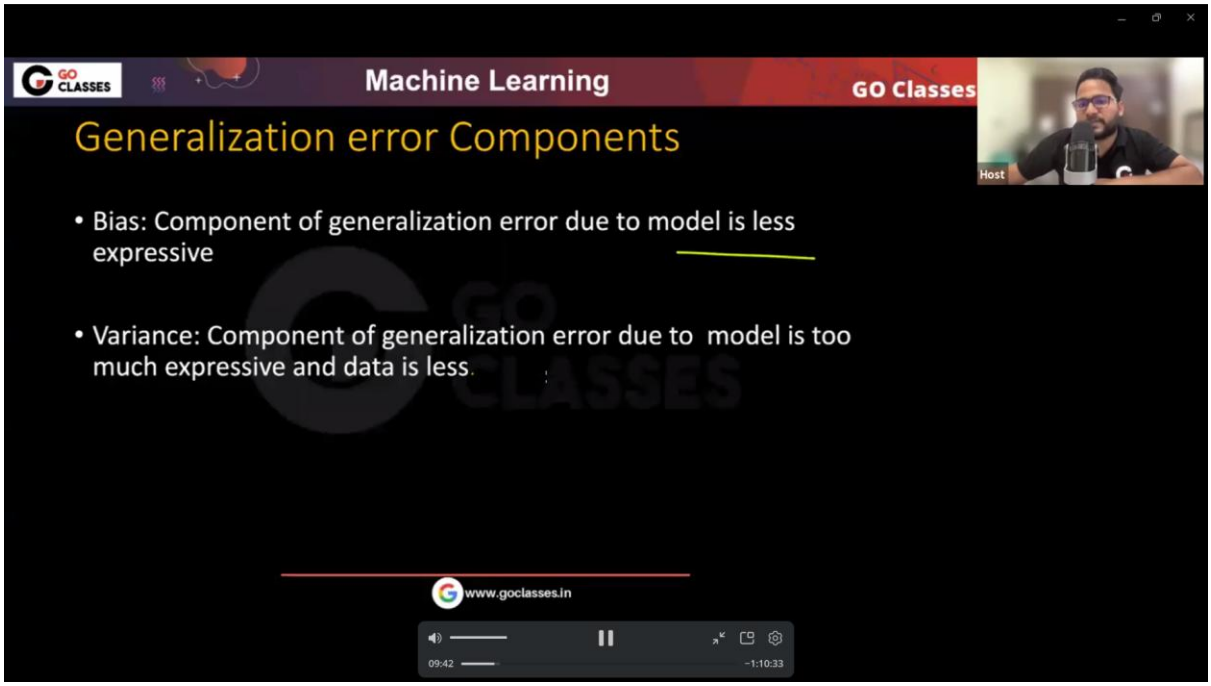

GO CLASSES
Machine Learning
GO Classes
Host
Guessing
underfitting
(simple model)
A
Mr. know it all
overfitting
B
Professor
Problem solving approach:
C
Just Right
02:02
-1:18:13

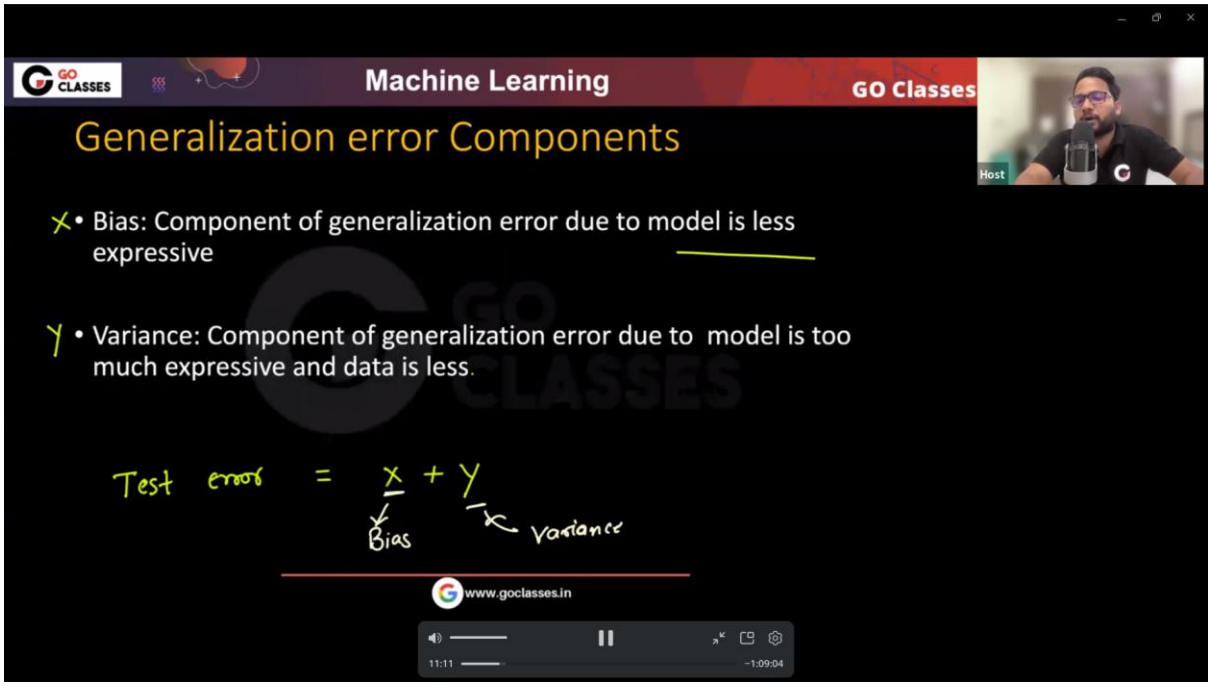

GO CLASSES
Machine Learning
GO Classes
Host
A
B
C
underfit
Just Right
overfit
www.goclasses.in
03:09
-1:17:06


GO CLASSES
Machine Learning
GO Classes
Host
www.goclasses.in
04:37
-1:15:38


GO CLASSES
Machine Learning
GO Classes
Under-fitting
Optimal-fitting
Over-fitting
Regression
Classification
www.goclasses.in
Host
05:03
-1:15:12

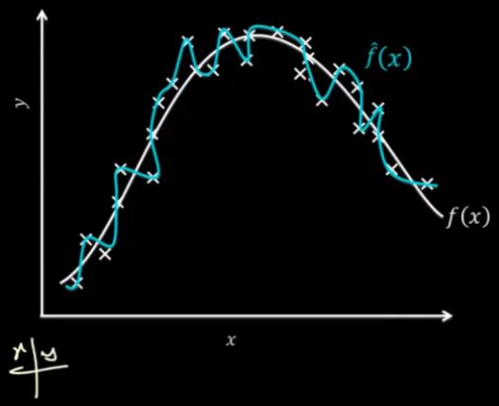

GO CLASSES
Machine Learning
GO Classes
Under-fitting
Optimal-fitting
Over-fitting
Regression
these both can not generalise but for seperate reasons
www.goclasses.in
Host





Bias→undefitting

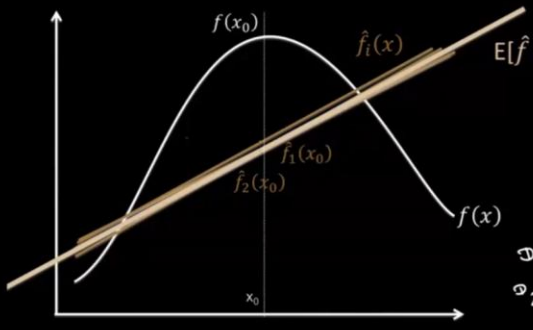
Variance→overfitting



- Let $f(x)$ be the true model (usually unknown)
- $MSE = \frac{1}{n}\Sigma(y_i - \hat{y}_i)^2$
- $\hat{y} = \hat{f}(x)$ — $\hat{f}$ could be simple or complex model

$\hat{f}(x)$ is complex model in this case.

14:18 −1:05:57

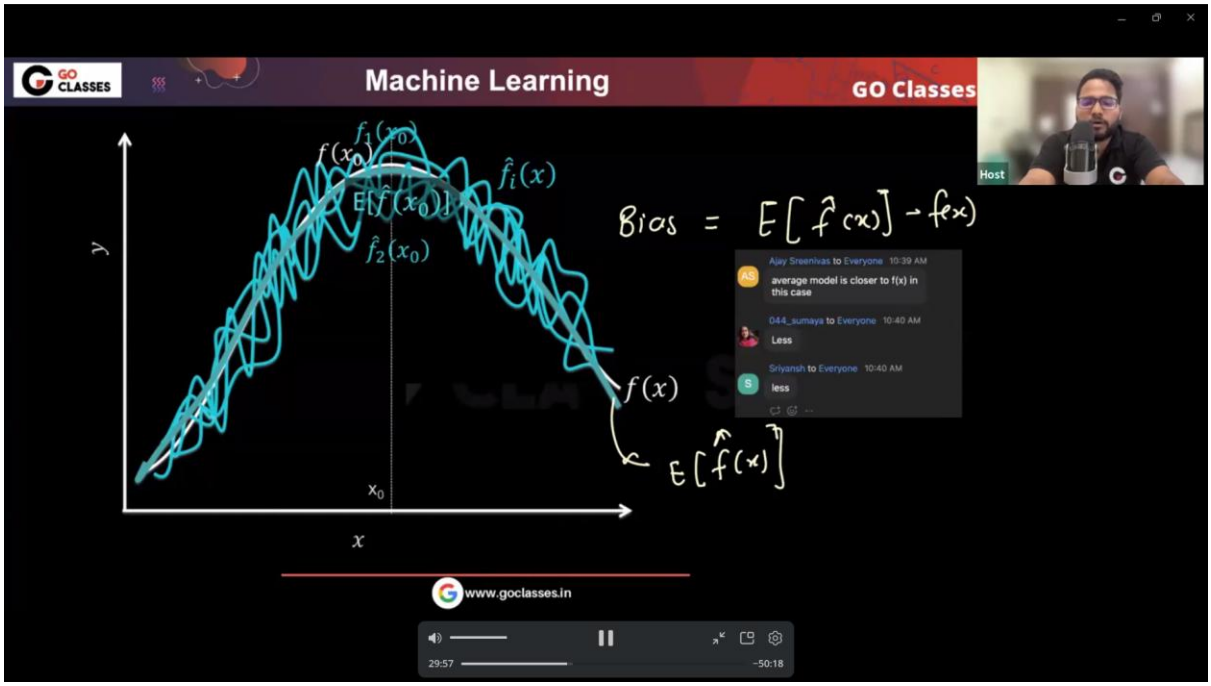


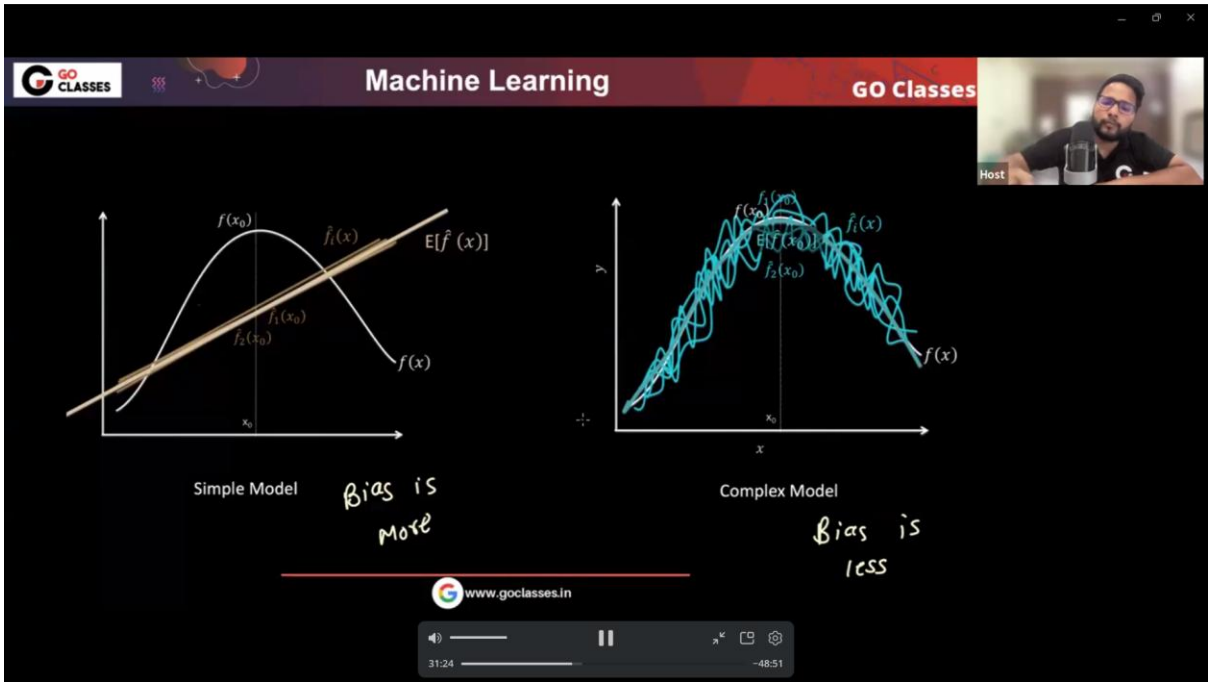
We can keep on taking different datasets and keep on getting new models.

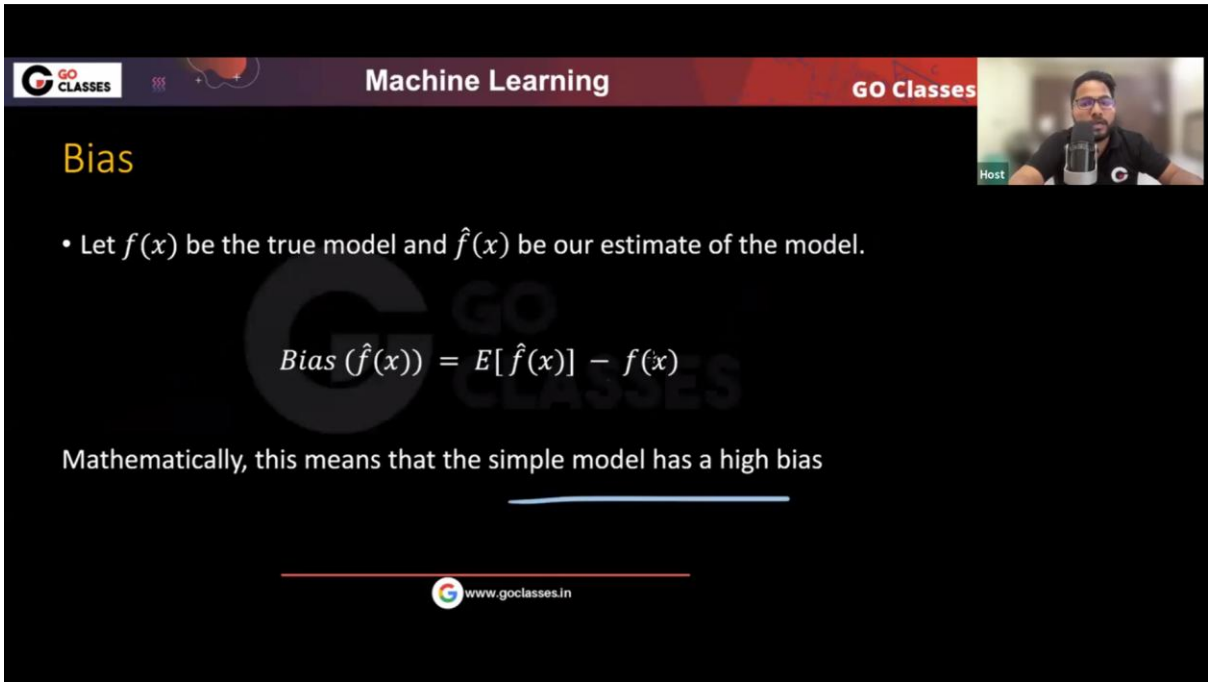
$$\text{Bias} = E[\hat{f}(x)] - f(x)$$

won't this term will be higher for simple model?

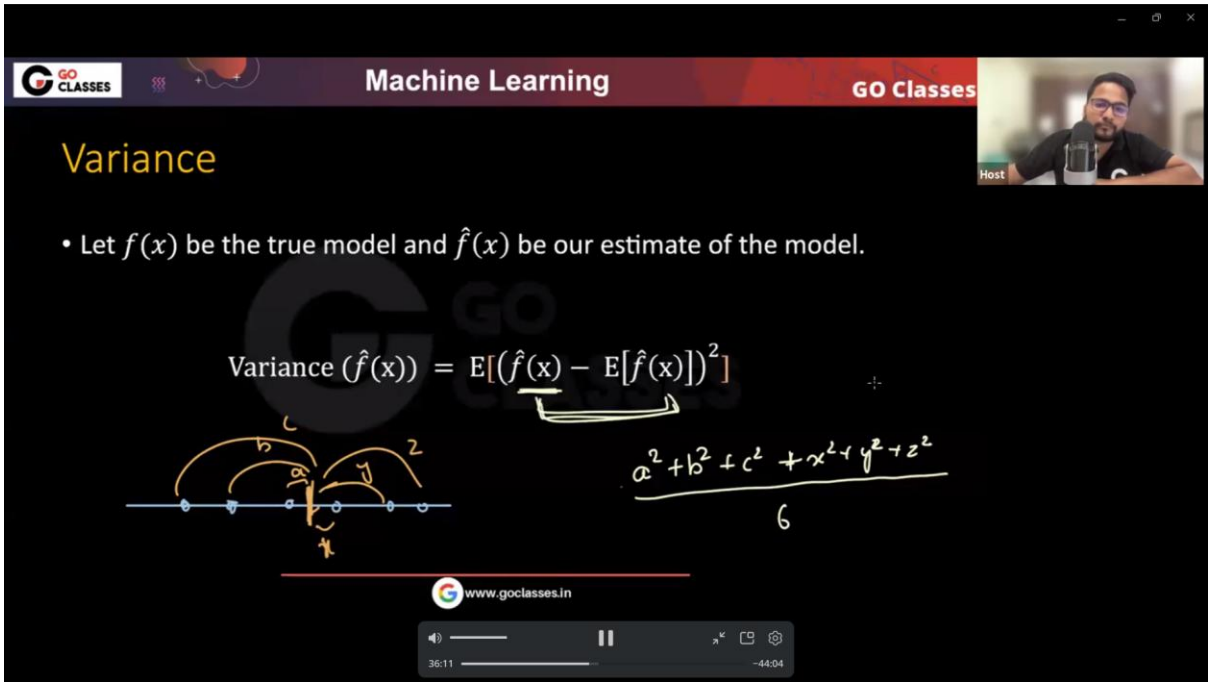
28:12 −52:03

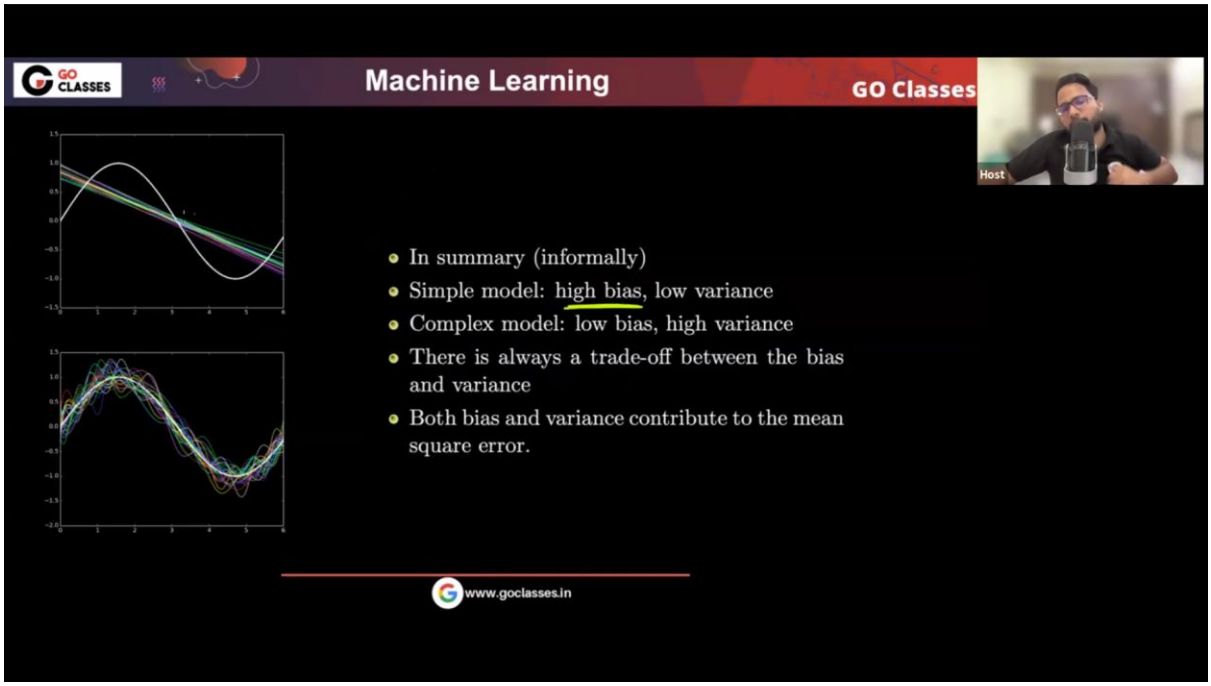

GO CLASSES
Machine Learning
GO Classes
Host
$f_1(x_0)$
$f(x_0)$
$\hat{f}_1(x)$
$E[\hat{f}(x_0)]$
$\hat{f}_2(x_0)$
$f(x)$
$E[\hat{f}(x)]$
x_0
y
x
Bias = $E[\hat{f}(x)] - f(x)$
Ajay Sreenivas to Everyone 10:39 AM
average model is closer to f(x) in this case
044_sumaya to Everyone 10:40 AM
Less
Srlyansh to Everyone 10:40 AM
less
www.goclasses.in
29:57
-50:18


GO CLASSES
Machine Learning
GO Classes
Host
$f(x_0)$
$\hat{f}_1(x)$
$E[\hat{f}(x)]$
$f_1(x_0)$
$\hat{f}_2(x_0)$
$f(x)$
x_0
Simple Model
Bias is more
Complex Model
Bias is less
www.goclasses.in
31:24
-48:51

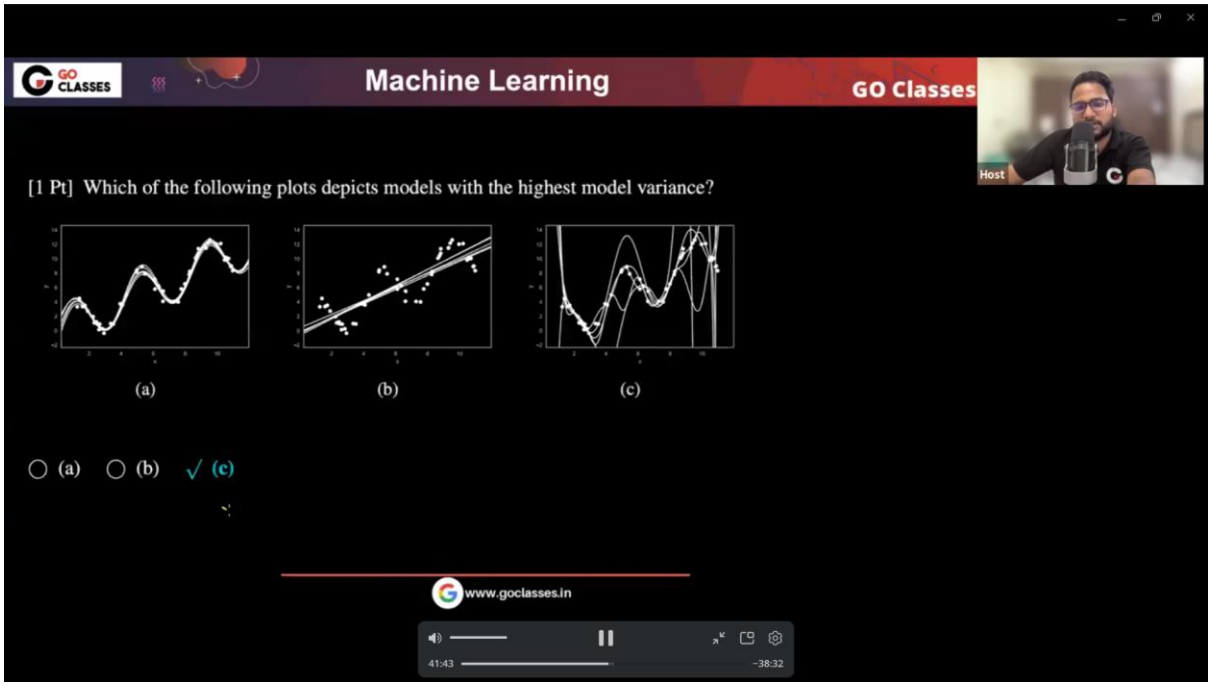

GO CLASSES
Machine Learning
GO Classes
Host
Bias
• Let $f(x)$ be the true model and $\hat{f}(x)$ be our estimate of the model.
$Bias\ (\hat{f}(x)) = E[\hat{f}(x)] - f(x)$
Mathematically, this means that the simple model has a high bias
www.goclasses.in

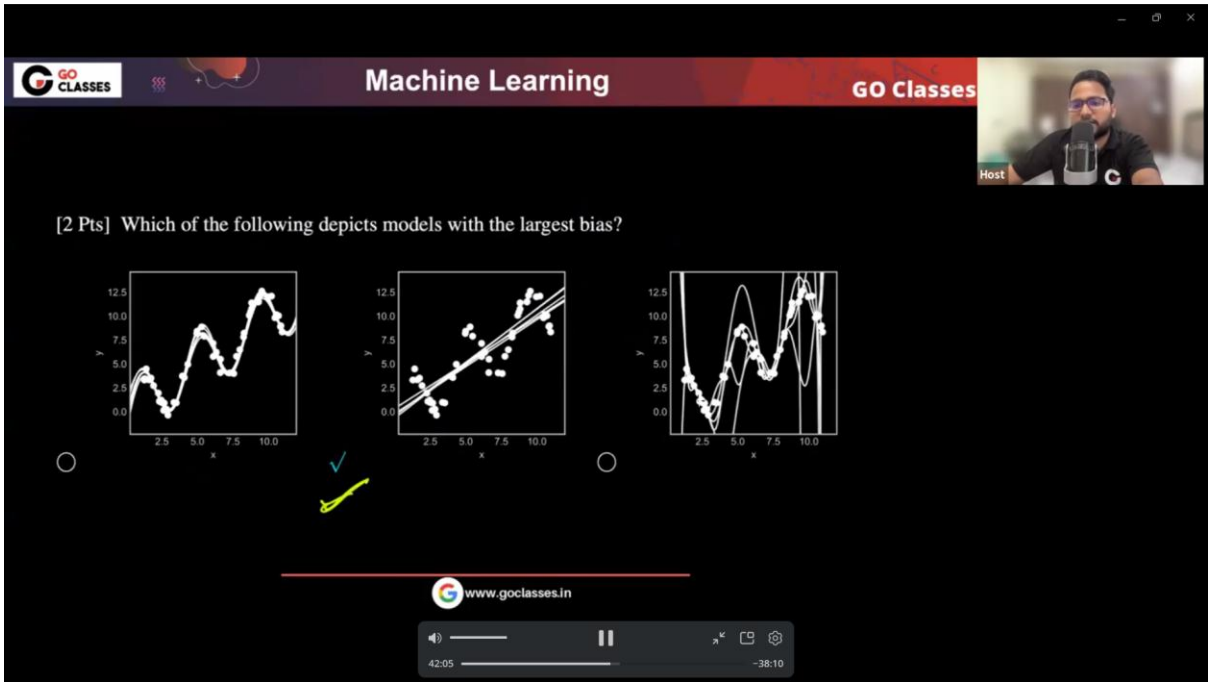
E→ estimated

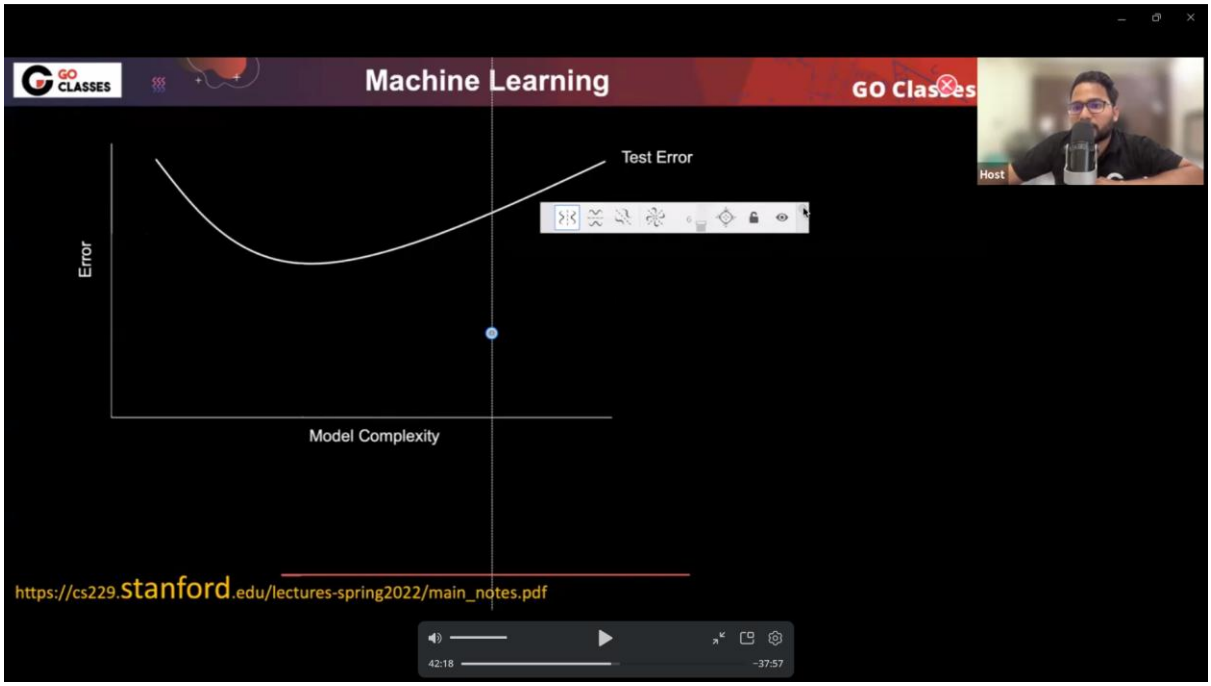

GO CLASSES
Machine Learning
GO Classes
Host
Variance
• Let $f(x)$ be the true model and $\hat{f}(x)$ be our estimate of the model.
Variance $(\hat{f}(x)) = E[(\hat{f}(x) - E[\hat{f}(x)])^2]$
$\frac{a^2+b^2+c^2+x^2+y^2+z^2}{6}$
www.goclasses.in
36:11
-44:04



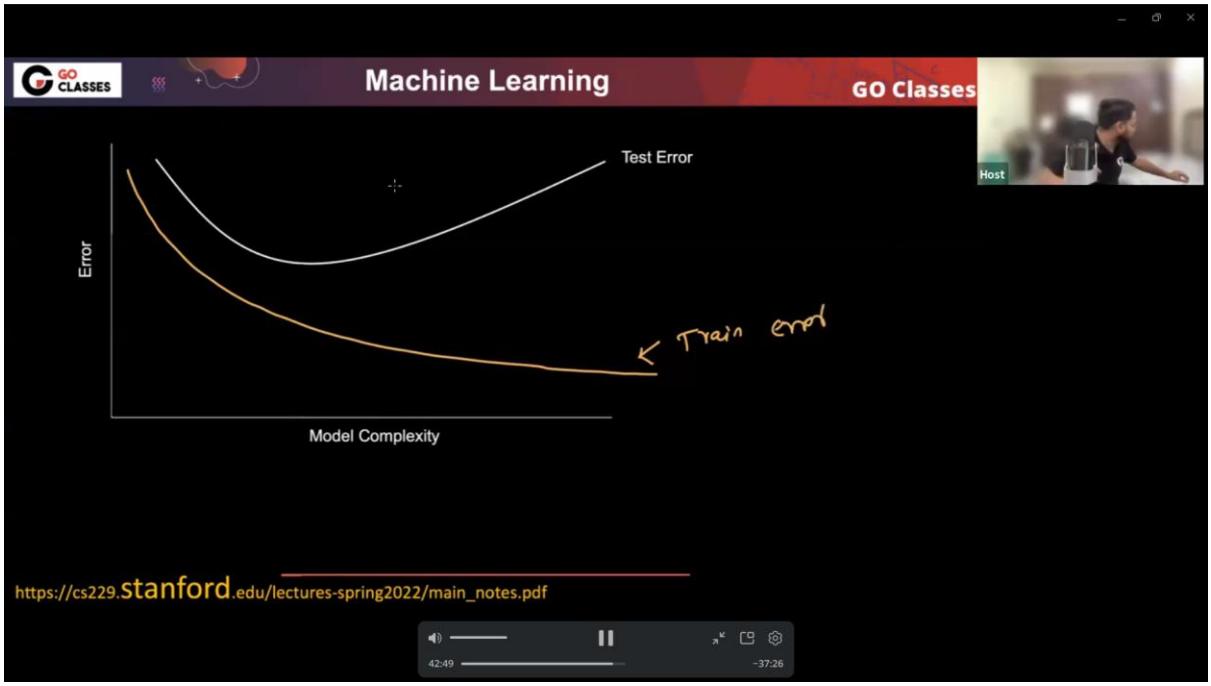
Bias→ how far it is from true function

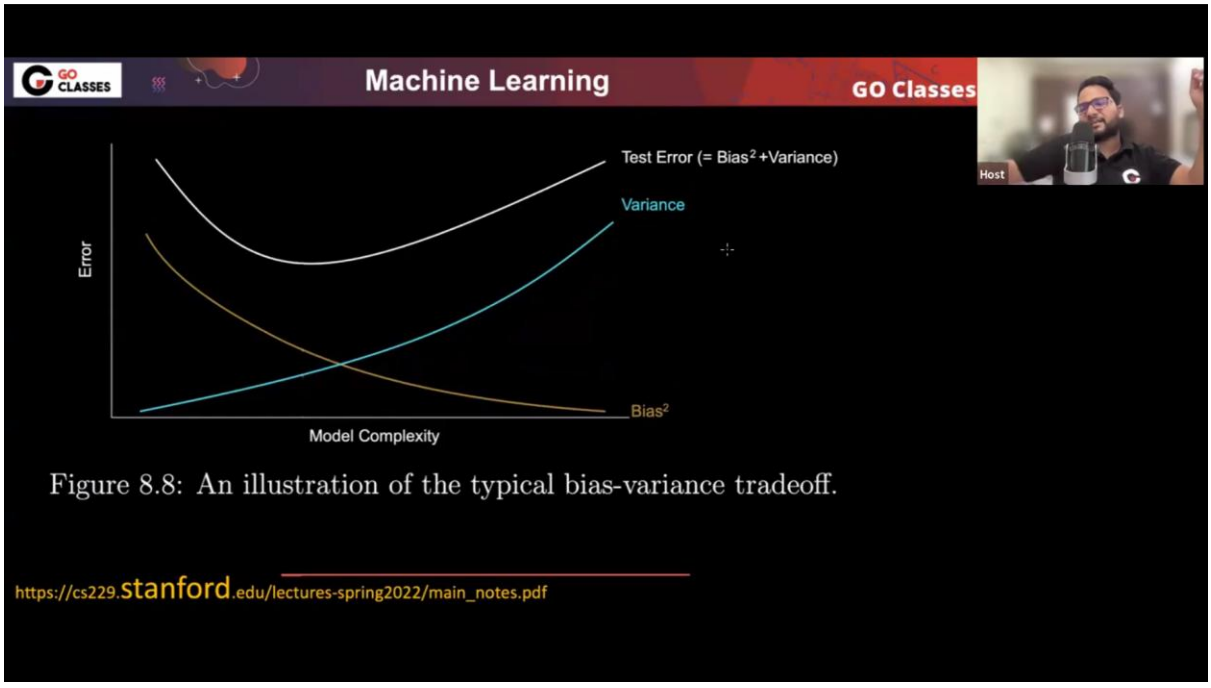




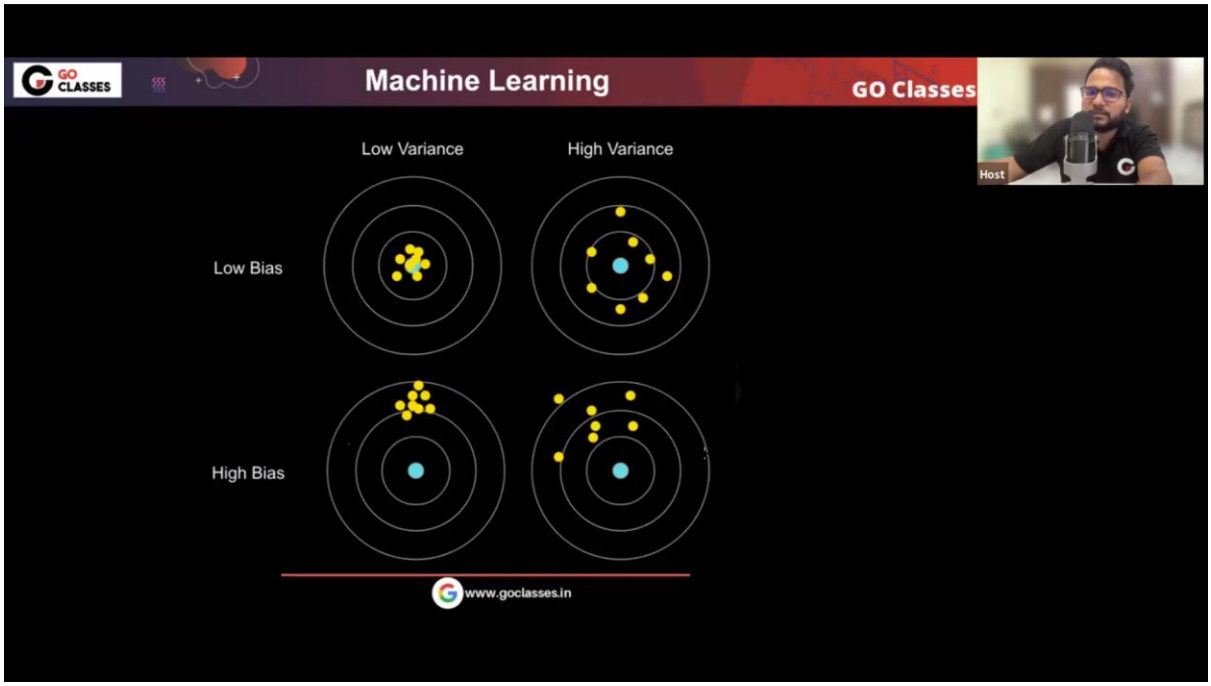


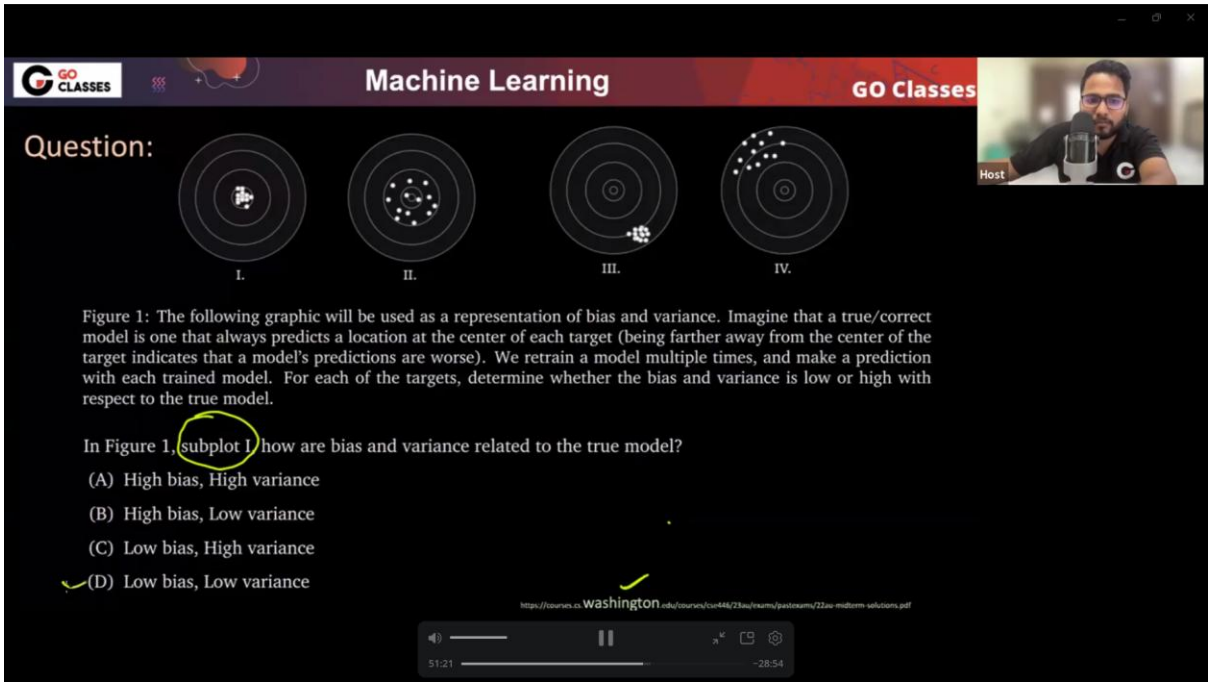
If we keep on increasing the complexity at some point the test error will also increase..





Predictions distance less→less variance


GO CLASSES
Machine Learning
GO Classes
Host
Low Variance
High Variance
Low Bias
High Bias
www.goclasses.in


GO CLASSES
Machine Learning
GO Classes
Host
Question:
I.
II.
III.
IV.
Figure 1: The following graphic will be used as a representation of bias and variance. Imagine that a true/correct model is one that always predicts a location at the center of each target (being farther away from the center of the target indicates that a model's predictions are worse). We retrain a model multiple times, and make a prediction with each trained model. For each of the targets, determine whether the bias and variance is low or high with respect to the true model.
In Figure 1, subplot I, how are bias and variance related to the true model?
(A) High bias, High variance
(B) High bias, Low variance
(C) Low bias, High variance
(D) Low bias, Low variance
https://courses.cs.washington.edu/courses/cse446/23au/exams/pastexams/22au-midterm-solutions.pdf
51:21
-28:54

GO CLASSES

Machine Learning

GO Classes

Question:

Figure 1: The following graphic will be used as a representation of bias and variance. Imagine that a true/correct model is one that always predicts a location at the center of each target (being farther away from the center of the target indicates that a model's predictions are worse). We retrain a model multiple times, and make a prediction with each trained model. For each of the targets, determine whether the bias and variance is low or high with respect to the true model.

In Figure 1, subplot II, how are bias and variance related to the true model?

(A) High bias, High variance

(B) High bias, Low variance

(C) Low bias, High variance

(D) Low bias, Low variance

https://courses.cs.washington.edu/courses/cse446/23au/exams/pastexams/22au-midterm-solutions.pdf

51:55 −28:20

GO CLASSES

Machine Learning

GO Classes

Question:

Figure 1: The following graphic will be used as a representation of bias and variance. Imagine that a true/correct model is one that always predicts a location at the center of each target (being farther away from the center of the target indicates that a model's predictions are worse). We retrain a model multiple times, and make a prediction with each trained model. For each of the targets, determine whether the bias and variance is low or high with respect to the true model.

In Figure 1, subplot II, how are bias and variance related to the true model?

(A) High bias, High variance

(B) High bias, Low variance

(C) Low bias, High variance

(D) Low bias, Low variance

https://courses.cs.washington.edu/courses/cse446/23au/exams/pastexams/22au-midterm-solutions.pdf

52:07 −28:08

GO CLASSES

Machine Learning

GO Classes

Question:

Figure 1: The following graphic will be used as a representation of bias and variance. Imagine that a true/correct model is one that always predicts a location at the center of each target (being farther away from the center of the target indicates that a model's predictions are worse). We retrain a model multiple times, and make a prediction with each trained model. For each of the targets, determine whether the bias and variance is low or high with respect to the true model.

In Figure 1, subplot II, how are bias and variance related to the true model?

(A) High bias, High variance

(B) High bias, Low variance

(C) Low bias, High variance

(D) Low bias, Low variance

https://courses.cs.washington.edu/courses/cse446/23au/exams/pastexams/22au-midterm-solutions.pdf

52:18 −27:57

Question:

[1 points] What best defines the relationship between a model's fit and its bias and variance

(a) A model with low bias and high variance is underfitting

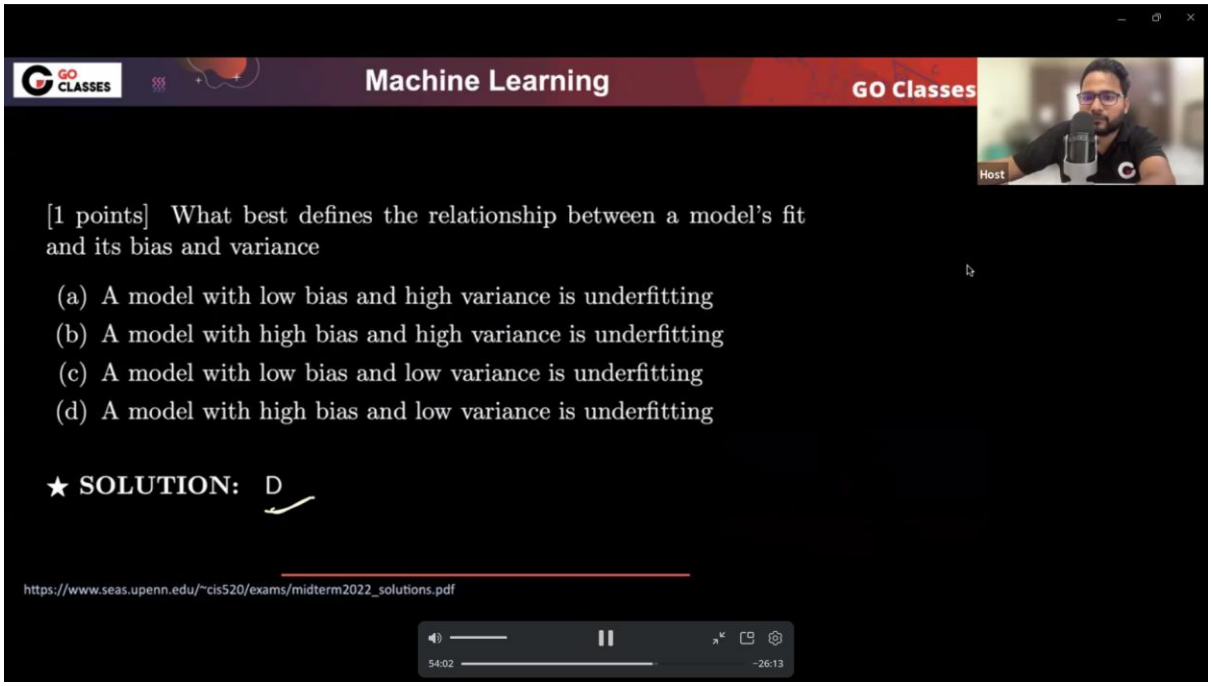
(b) A model with high bias and high variance is underfitting

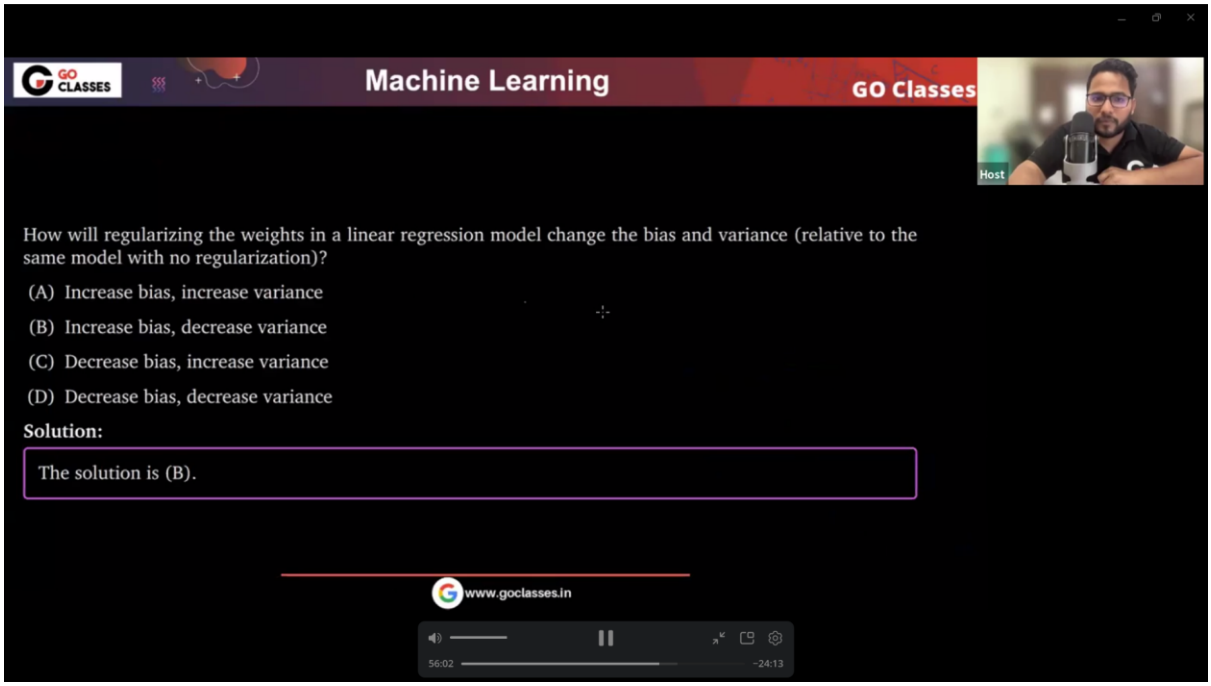
(c) A model with low bias and low variance is underfitting

(d) A model with high bias and low variance is underfitting

https://www.seas.upenn.edu/~cis520/exams/midterm2022_solutions.pdf

53:53 −26:22


GO CLASSES
Machine Learning
GO Classes
Host
[1 points] What best defines the relationship between a model's fit and its bias and variance
(a) A model with low bias and high variance is underfitting
(b) A model with high bias and high variance is underfitting
(c) A model with low bias and low variance is underfitting
(d) A model with high bias and low variance is underfitting
★ SOLUTION: D
https://www.seas.upenn.edu/~cis520/exams/midterm2022_solutions.pdf
54:02
-26:13


GO CLASSES
Machine Learning
GO Classes
Host
How will regularizing the weights in a linear regression model change the bias and variance (relative to the same model with no regularization)?
(A) Increase bias, increase variance
(B) Increase bias, decrease variance
(C) Decrease bias, increase variance
(D) Decrease bias, decrease variance
Solution:
The solution is (B).
www.goclasses.in
56:02
-24:13

Question:

How will regularizing the weights in a linear regression model change the bias and variance (relative to the same model with no regularization)?

(A) Increase bias, increase variance

(B) Increase bias, decrease variance

(C) Decrease bias, increase variance

(D) Decrease bias, decrease variance

Regularization
⇓
reduce overfitting
⇓
reduce variance

https://courses.cs.washington.edu/courses/cse446/23wi/assignments/midterm/midterm-solutions.pdf

(1 point) Peanut wants to train a model to accurately classify different types of animals from images. After training and testing his model, he observes that the model has high training error and high test error. What can we most confidently say about the bias/variance characteristics of Peanut's model?

A. High bias. → model is simple (because high training error is given)

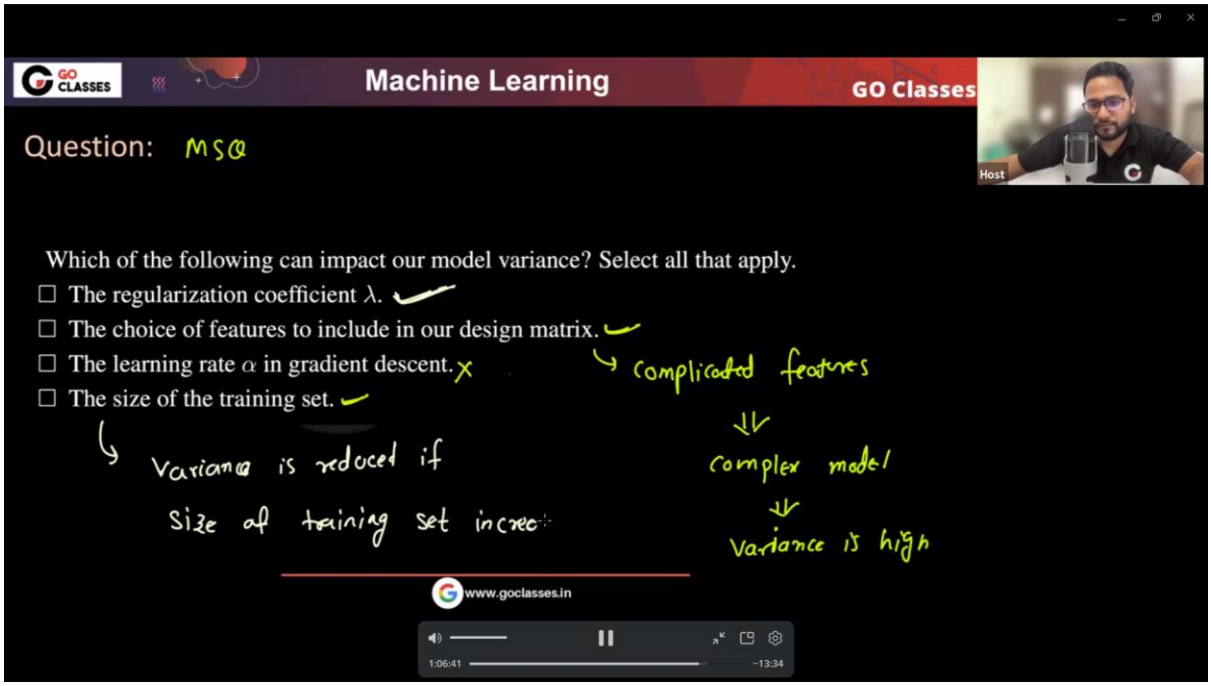
B. Low bias.

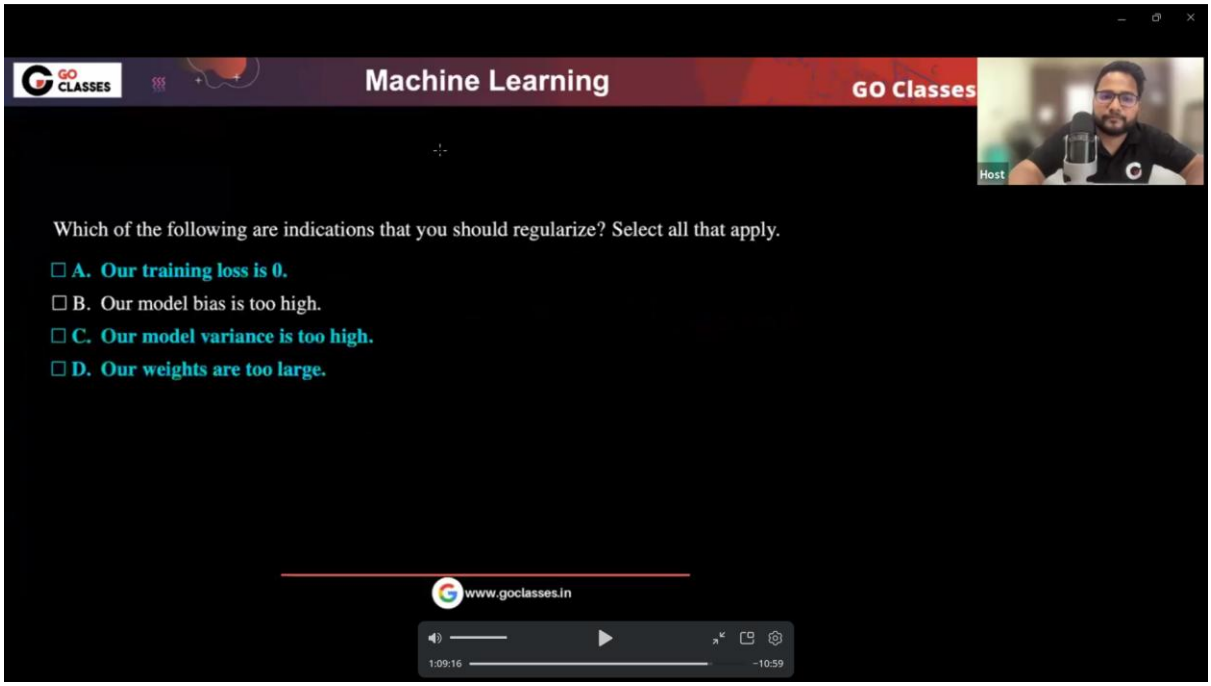
C. High variance.

D. Low variance.

Solution: A. High bias because the model has high training error.

High training error ⇒ Simple model
⇒ High bias
Low variance


GO CLASSES
Machine Learning
GO Classes
Host
Question: MSQ
Which of the following can impact our model variance? Select all that apply.
☐ The regularization coefficient λ.
☐ The choice of features to include in our design matrix.
☐ The learning rate α in gradient descent.
☐ The size of the training set.
complicated features
complex model
Variance is high
Variance is reduced if
Size of training set increa
www.goclasses.in
1:06:41
-13:34


GO CLASSES
Machine Learning
GO Classes
Host
Which of the following are indications that you should regularize? Select all that apply.
☐ A. Our training loss is 0.
☐ B. Our model bias is too high.
☐ C. Our model variance is too high.
☐ D. Our weights are too large.
www.goclasses.in
1:09:16
-10:59

Reducing the regularization of a model would typically . . .

(a) Decrease its bias and increase its variance

(b) Decrease its bias and decrease its variance

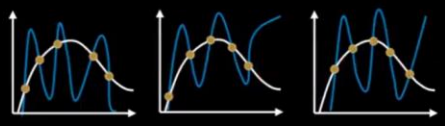
(c) Increase its bias and decrease its variance

(d) Increase its bias and increase its variance

Correct answers: (a)

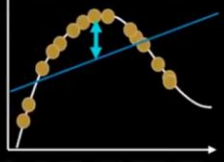
Variance

Regardless of what the true function is, how much does our prediction change with dataset?



Bias^2

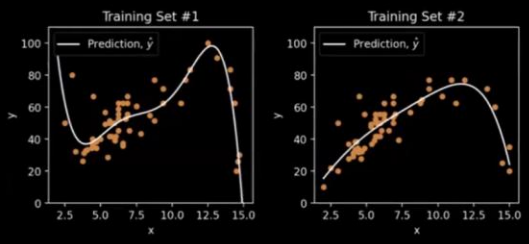
This error doesn't go away no matter how much data we have!



Variance reduces if we increase data

$\text{Bias} = E(\hat{f}(x)) - f(x)$

Consider a quintic (degree 5) model $\hat{y} = \hat{q}_0 + \hat{q}_1 x + ... + \hat{q}_5 x^5$, where $\hat{q}_j$ is the MSE parameter estimate for feature j of this quintic model. You first fit this model to Training Set #1; then, you fit this model again to Training Set #2.

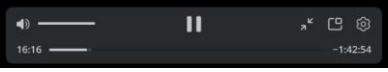


(i) The primary contributor to this model's risk is: ○ High bias ○ **High variance**

(ii) This contributor to model risk can be addressed by the following (select all that apply):

- ☐ **Reducing model complexity**
- ☐ Adding additional features to the model
- ☐ Reducing the regularization hyperparameter λ
- ☐ **Increasing the regularization hyperparameter λ**

https://ds100.org/sp24/resources/assets/exams/sp23/sp23_final_sol.pdf



(b) [1 Pt] Suppose we control the complexity of the linear models using a Ridge penalty term $\lambda \sum \beta_i^2$. Which of the following is true?

- ○ The left side of the graph represents small λ.
- ○ **The right side of the graph represents small λ.**

Solution: A smaller λ value means higher model complexity. Remember that a zero λ value means a model with no regularization.

(c) [3 Pts] Which of the following can impact our model variance? Select all that apply.

- ☐ **The regularization coefficient λ.**
- ☐ **The choice of features to include in our design matrix.**
- ☐ The learning rate α in gradient descent.
- ☐ **The size of the training set.**