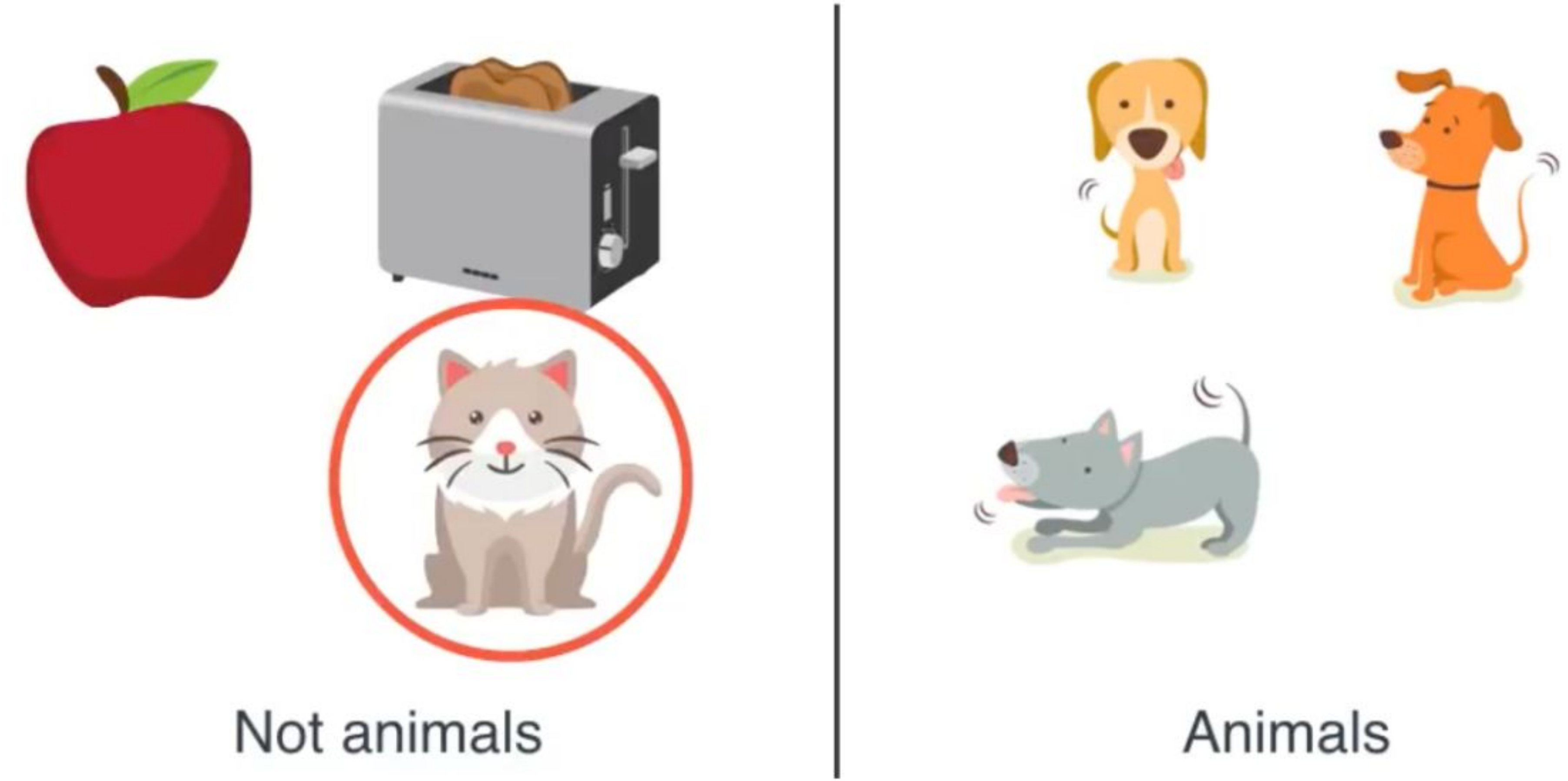
Bias Variance

Prepared By: **Dr.Mydhili K Nair, Professor, ISE Dept, RIT**
For: **Machine Learning Class**
Target Audience: **Sem 6 Students**

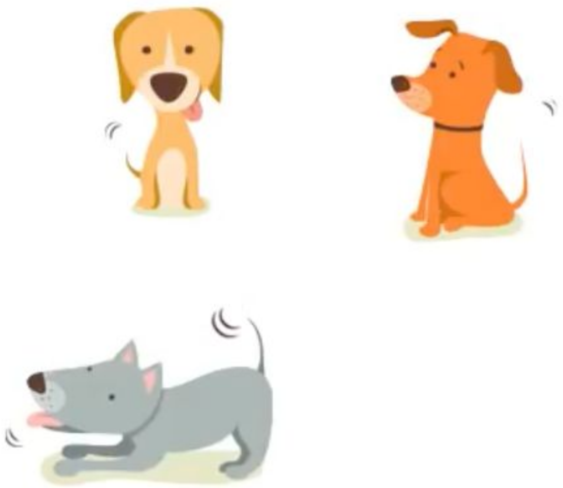
Bias


Not animals
Animals

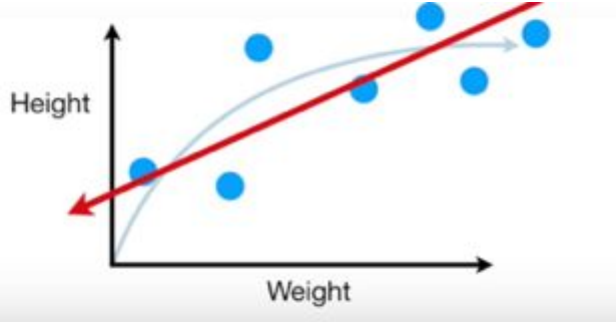
Our Classification Model is too specific

Anything but dogs that are wagging their tail

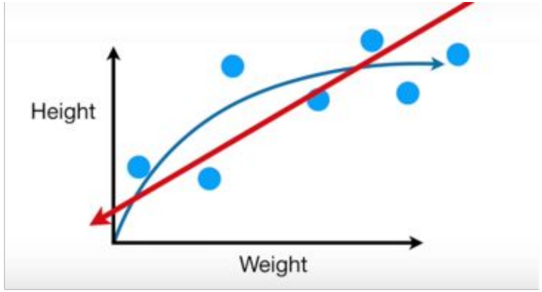
Too specific

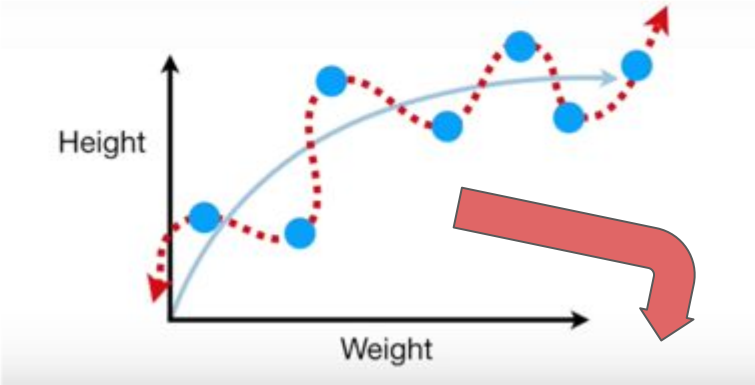


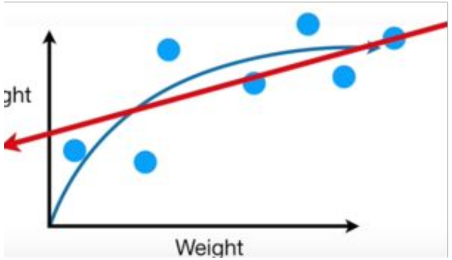
Dogs that are wagging their tail



Linear Model: How much ever you vary the "prediction regression line" it will not fit the curve and there will always be a **bias** between actual value and predicted value.



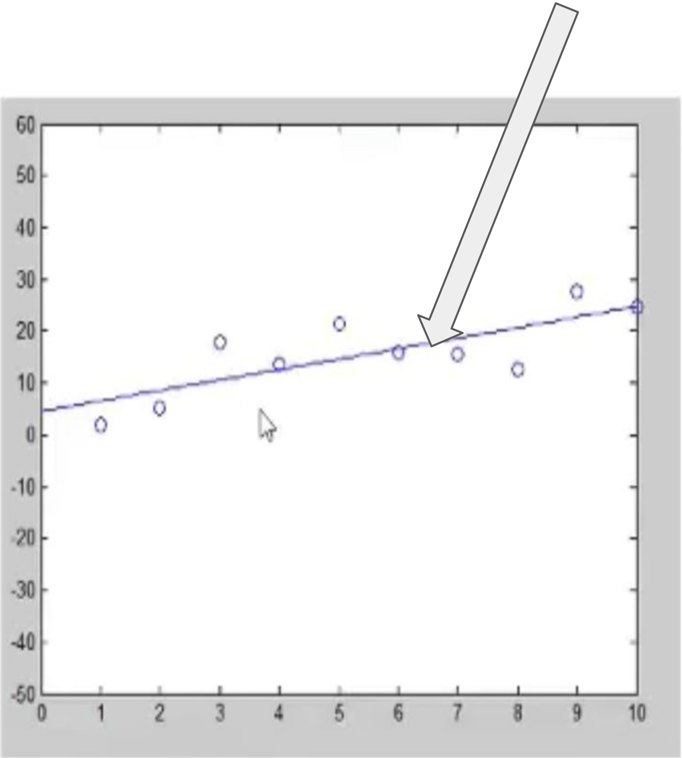




Bias error is completely eliminated. This non-linear model fits the data points perfectly. Zero Bias.

Model 1: $y = b_0 + b_1 x$

Linear Model

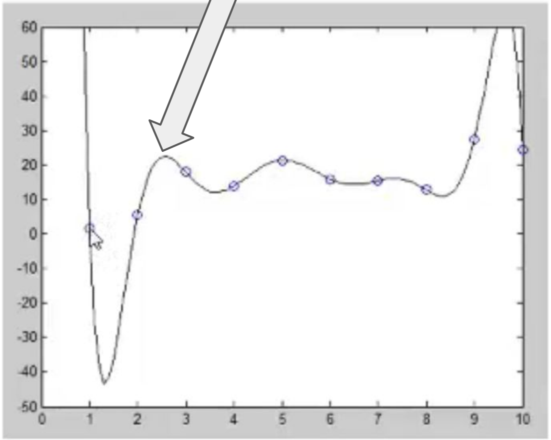


x	y
1	1.7
2	5.3
3	18.0
4	13.8
5	21.4
6	15.9
7	15.5
8	12.7
9	27.5
10	24.6

Model 2: $y = b_0 + b_1 x + b_2 x^2 + b_3 x^3 + \cdots + b_9 x^9$

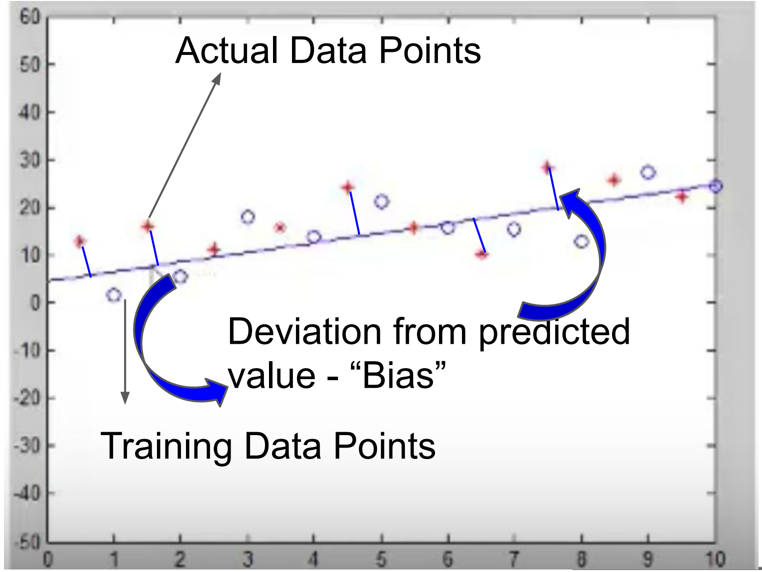
x	x^2		x^9	y
1	1	...	1	1.7
2	4		262144	5.3
3	9		4E+08	18.0
	16		7E+10	13.8
	25		4E+12	21.4
	36		1E+14	15.9
	49		2E+15	15.5
	64		2E+16	12.7
	81		2E+17	27.5
	100		1E+18	24.6

9th Order Polynomial Model



$$\text{Model 1: } y = b_0 + b_1 x$$

Our Classification Model is too simple - HIGH BIAS

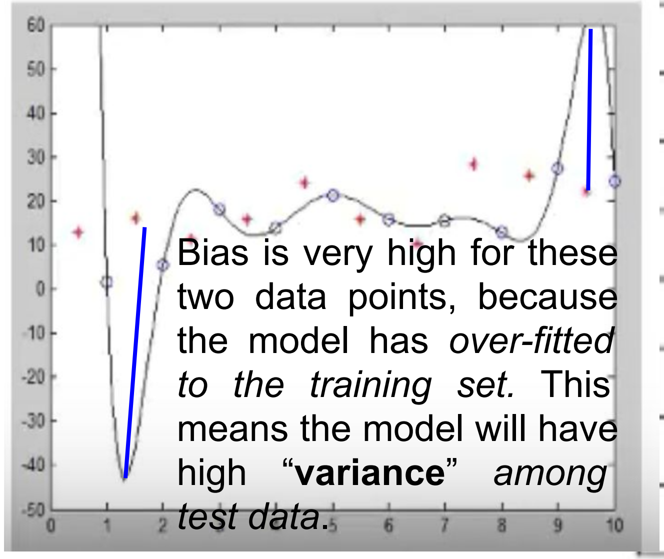


x	y
1	1.7
2	5.3
3	18.0
4	13.8
5	21.4
6	15.9
7	15.5
8	12.7
9	27.5
10	24.6

Model 2: $y = b_0 + b_1x + b_2x^2 + b_3x^3 + \cdots + b_9x^9$

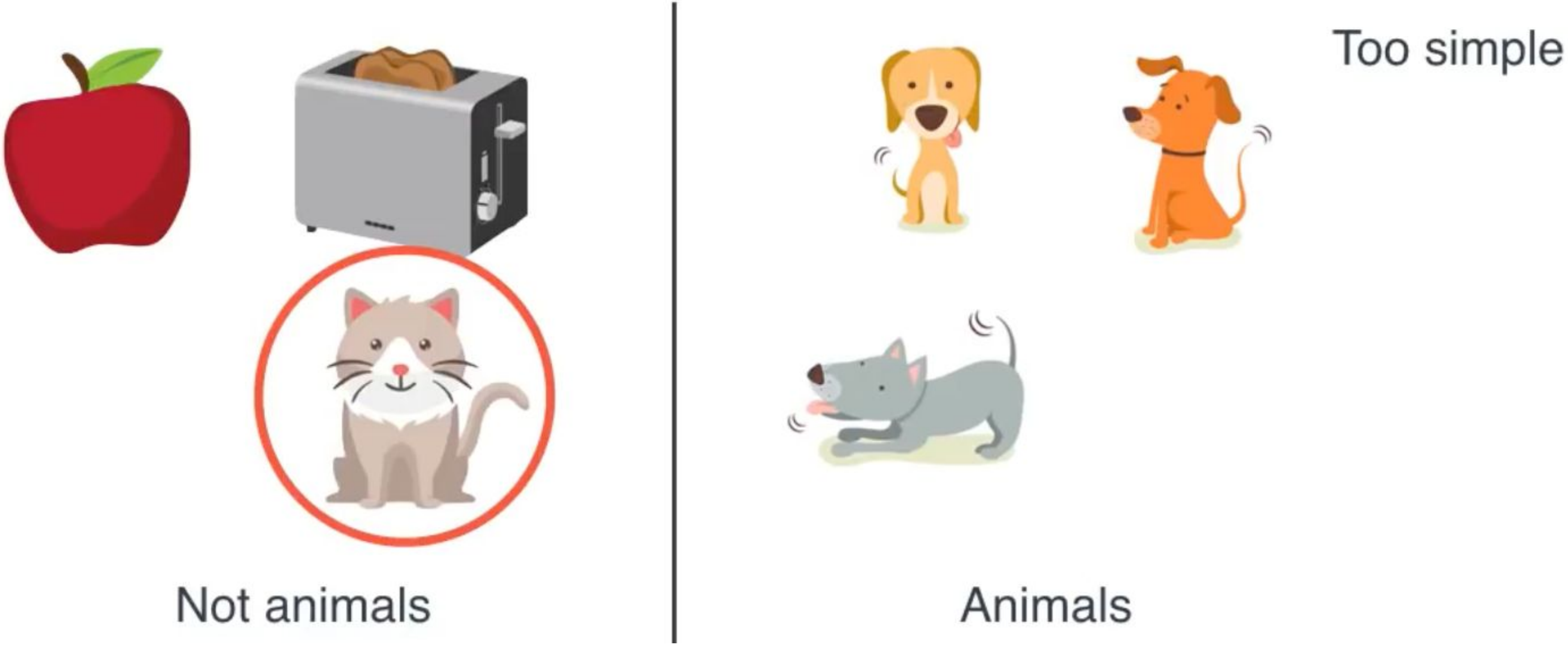
x	x^2		x^9	y
1	1	...	1	1.7
2	4		262144	5.3
3	9		4E+08	18.0
4	16		7E+10	13.8
5	25		4E+12	21.4
6	36		1E+14	15.9
7	49		2E+15	15.5
8	64		2E+16	12.7
9	81		2E+17	27.5
10	100		1E+18	24.6

Our Classification Model is too specific - HIGH VARIANCE



Bias is very high for these two data points, because the model has *over-fitted to the training set.* This means the model will have high “**variance**” *among test data.*

Error due to bias (underfitting)

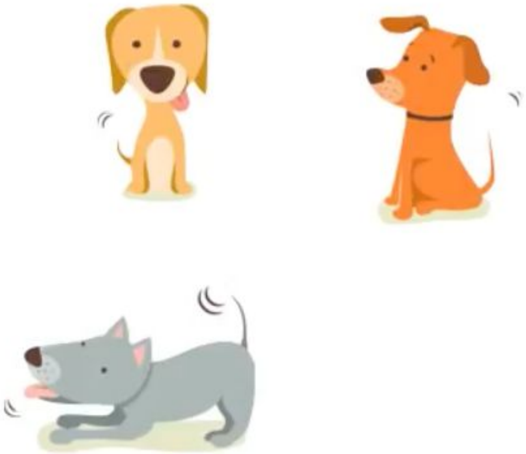


Prepared By: Dr.Mydhili K Nair, Prof, ISE Dept, Ramaiah Institute of Technology, Bengaluru

Error due to variance (overfitting)

Anything but dogs that are wagging their tail

Too specific

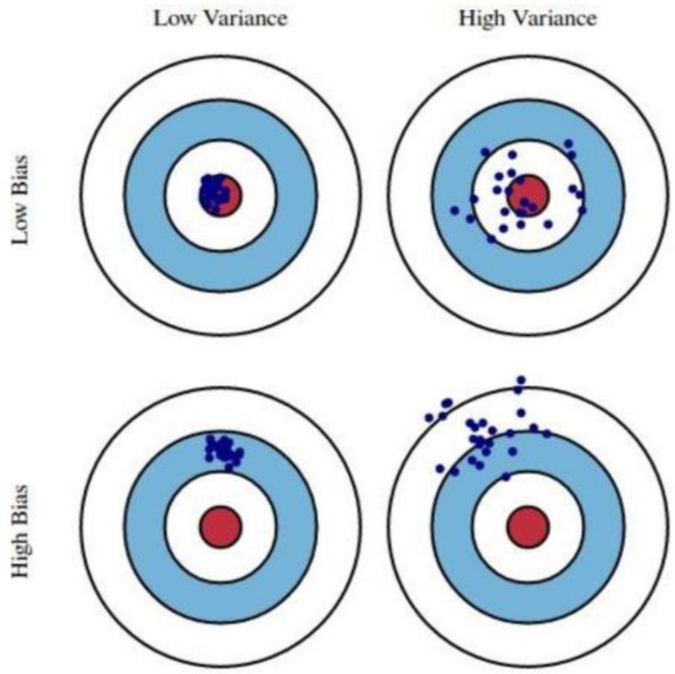


Dogs that are wagging their tail

Trade-off between Bias-variance

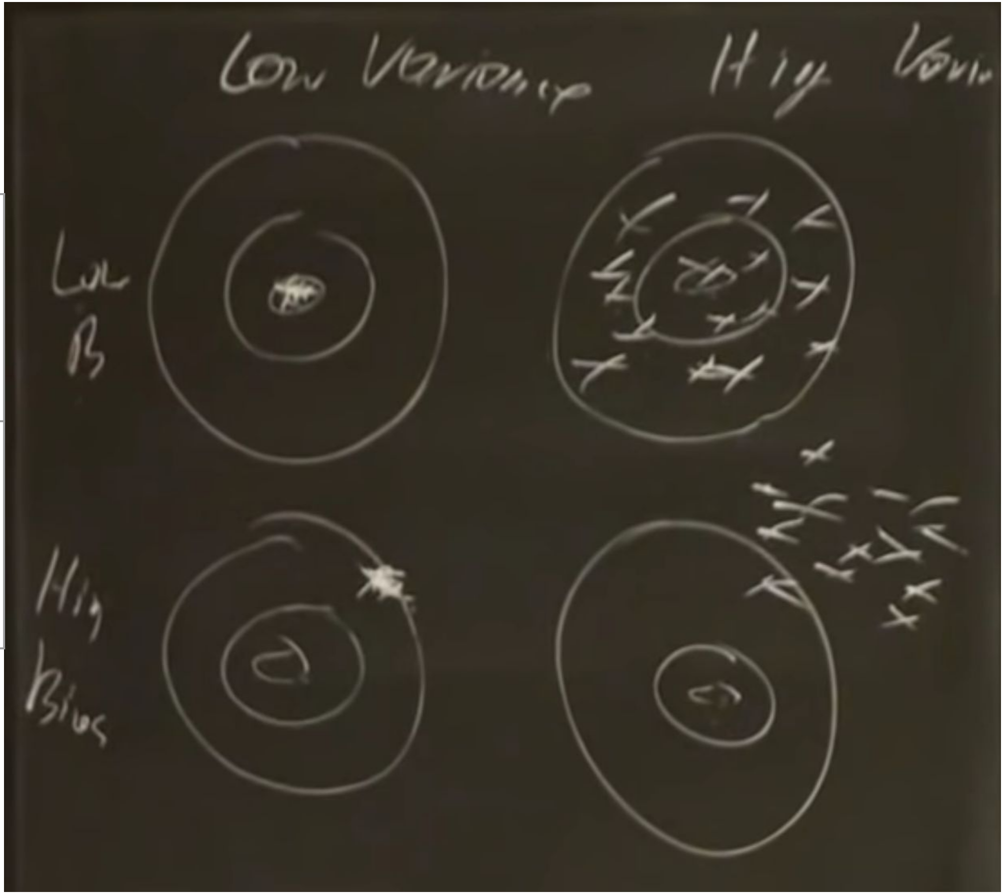
"Bias is the algorithm's tendency to consistently learn the wrong thing by not taking into account all the information in the data (underfitting)."

"Variance is the algorithm's tendency to learn random things irrespective of the real signal by fitting highly flexible models that follow the error/noise in the data too closely (overfitting)."



Bull's eye diagram

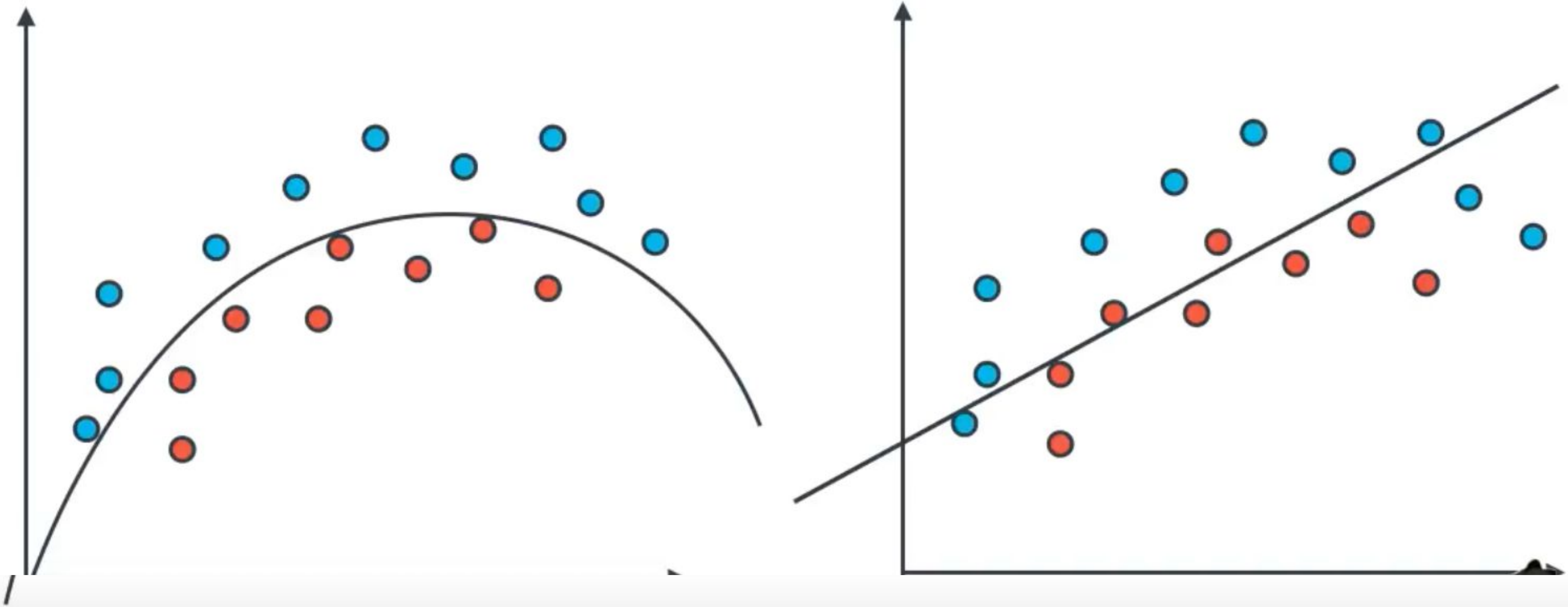
	Low Variance	High Variance
Low Bias		
High Bias		



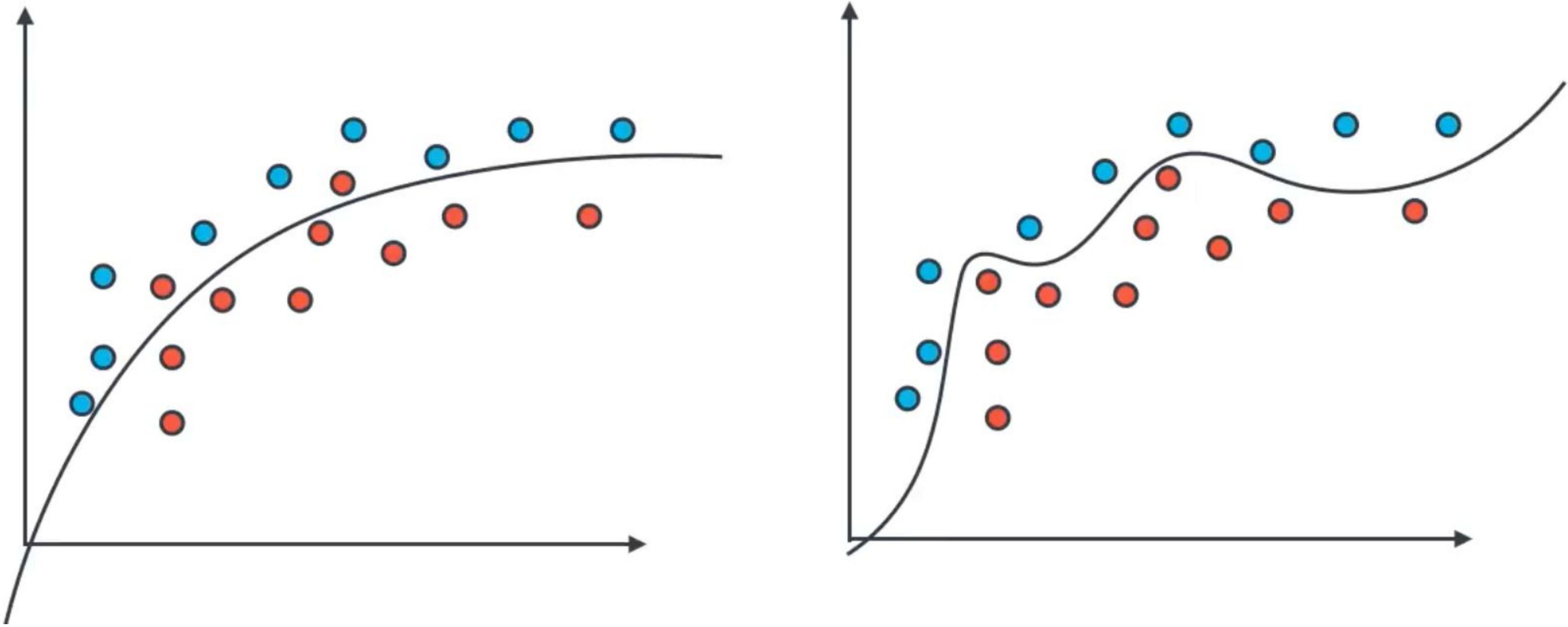
Every Model aims at *Low Bias & Low Variance*

Bias - Variance Tradeoff
Overfitting - Underfitting

Error due to bias (underfitting)



Error due to variance (overfitting)



Overfitting (aka variance):

A model is said to be overfit if it is over trained on the data such that, it even learns the noise from it. An overfit model learns each and every example so perfectly that it misclassifies an unseen/new example. For a model that's overfit, we have a perfect/close to perfect training set score while a poor test/validation score.

Reasons behind overfitting:

1. Using a complex model for a simple problem which picks up the noise from the data. Example: Fitting a neural network to the Iris dataset.
2. Small datasets, as the training set may not be a right representation of the universe.

Source: https://towardsdatascience.com/learning-curve-to-identify-overfitting-underfitting-problems-133177f38df5

Underfitting (aka bias):

A model is said to be underfit if it is unable to learn the patterns in the data properly. An underfit model doesn't fully learn each and every example in the dataset. In such cases, we see a low score on both the training set and test/validation set.

Reasons behind underfitting:

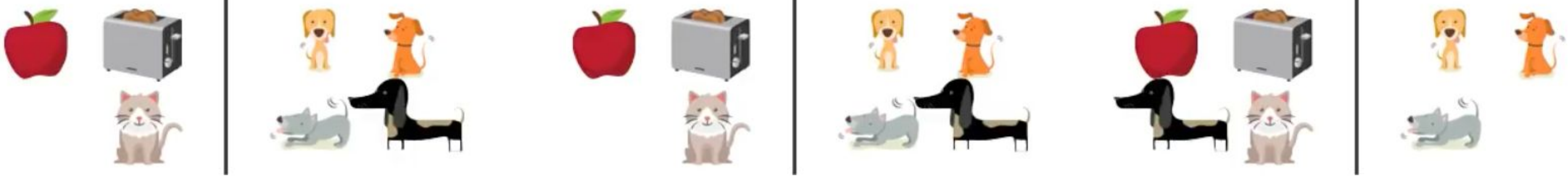
1. Using a simple model for a complex problem which doesn't learn all the patterns in the data. Example: Using a logistic regression for image classification

2. The underlying data has no inherent pattern. Example, trying to predict a student's marks with his father's weight.

Source: https://towardsdatascience.com/learning-curve-to-identify-overfitting-underfitting-problems-133177f38df5

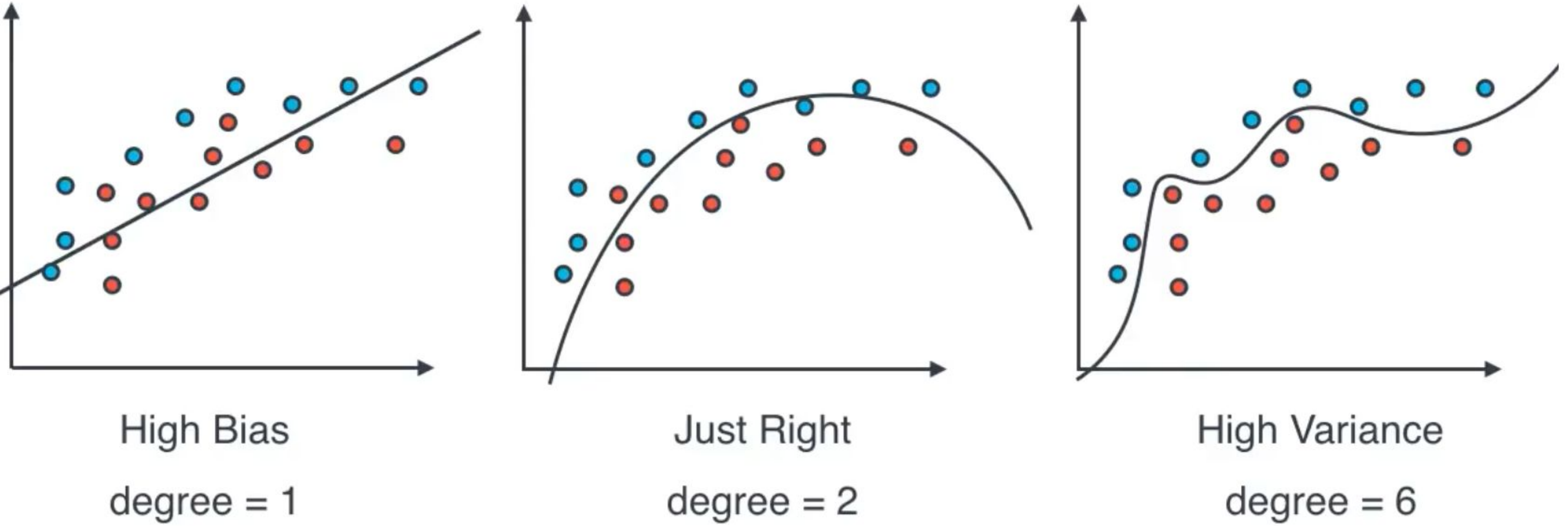
Tradeoff

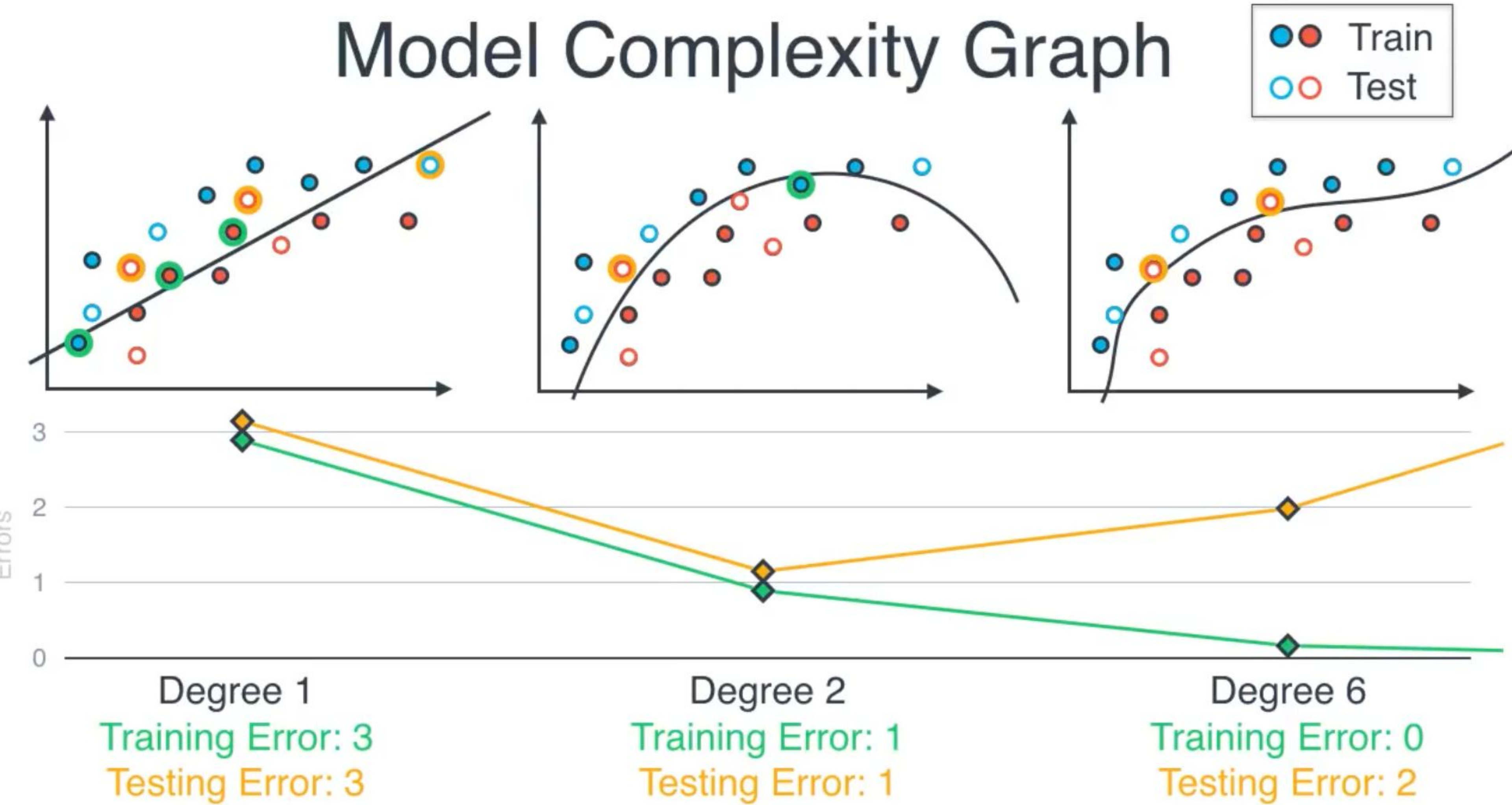
High bias (Underfitting)	Just Right	High variance (Overfitting)
Not animals | Animals	Not dogs | Dogs	Not dogs who wag their tail | Dogs who wag their tail



High bias (Underfitting)	Just Right	High variance (Overfitting)
Bad on Training set	Good on Training set	Great on Training set
Bad on Testing set	Good on Testing set	Bad on Testing set

Model Complexity Graph

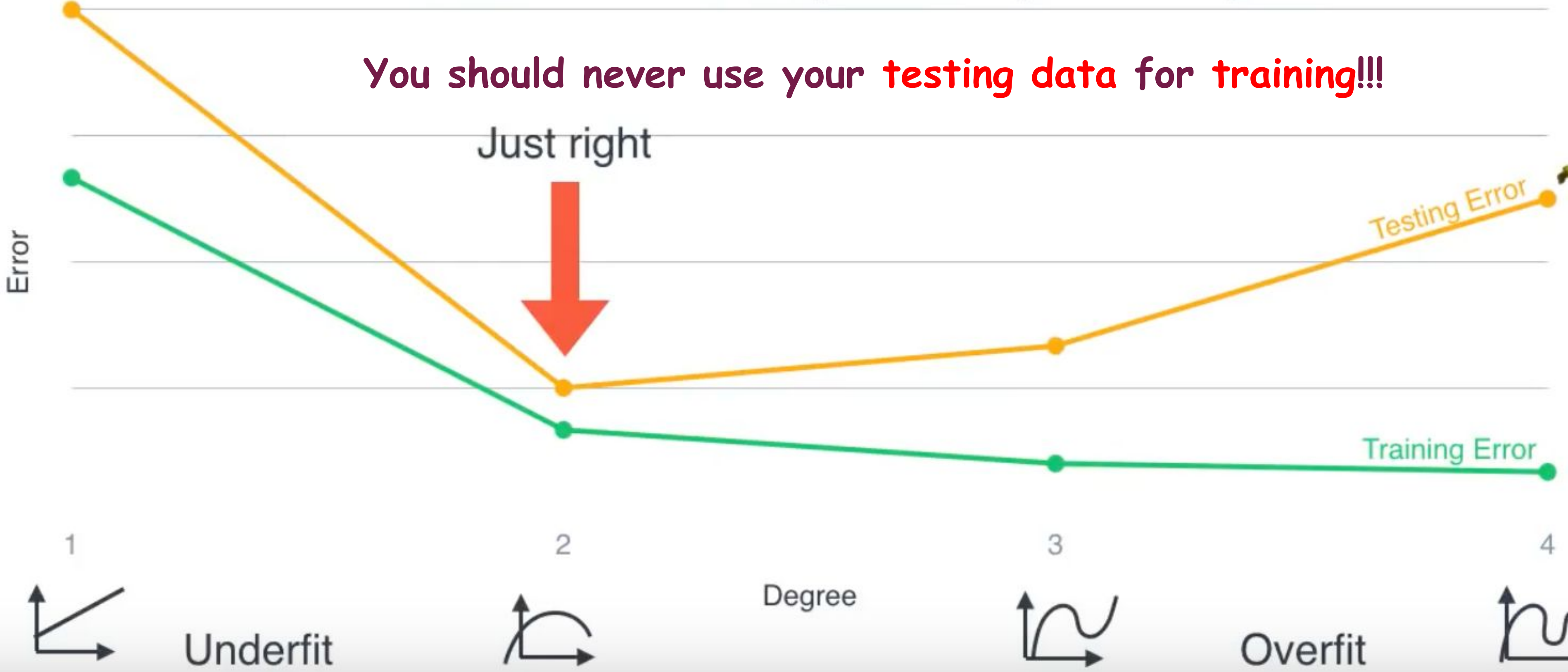




Model Complexity Graph
Train
Test
3
2
1
0
Degree 1
Training Error: 3
Testing Error: 3
Degree 2
Training Error: 1
Testing Error: 1
Degree 6
Training Error: 0
Testing Error: 2

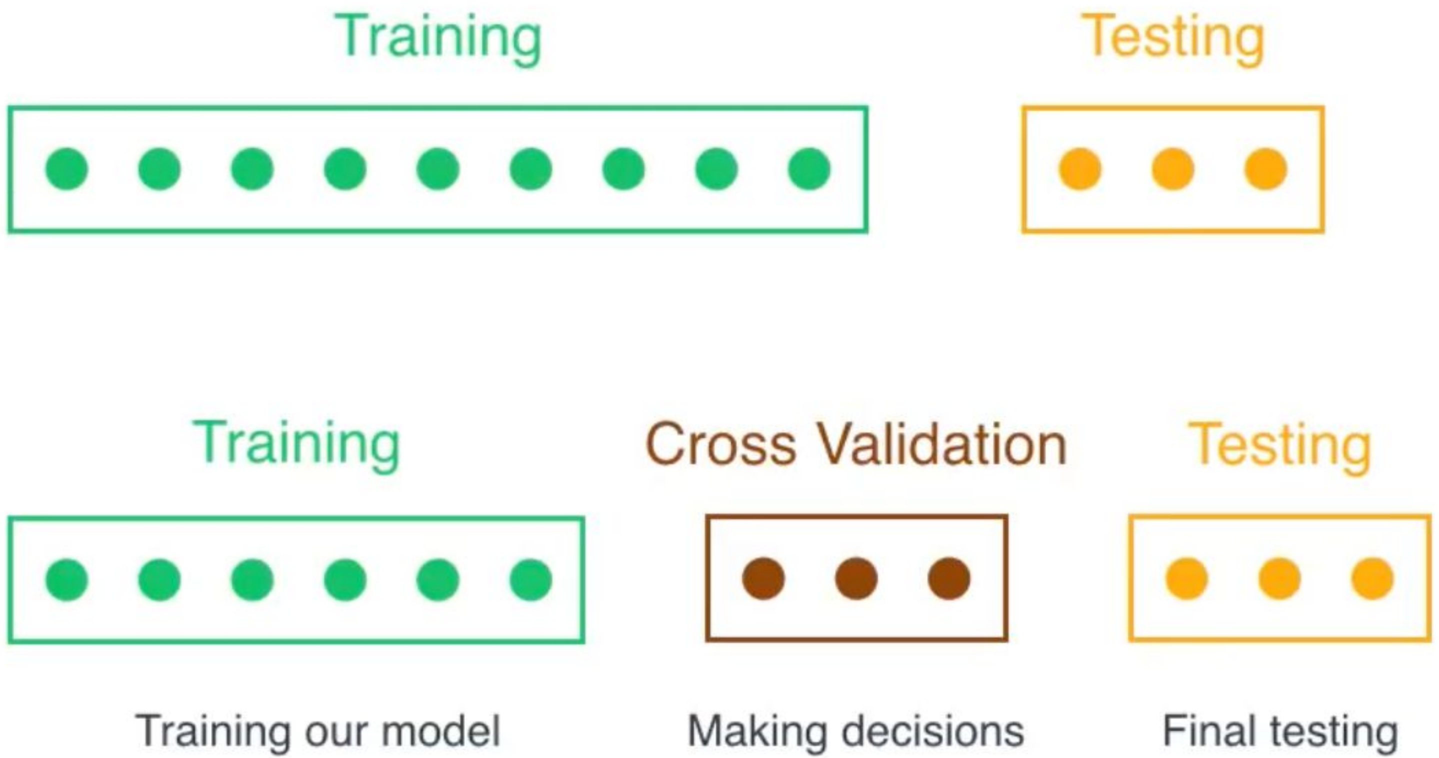
Model Complexity Graph

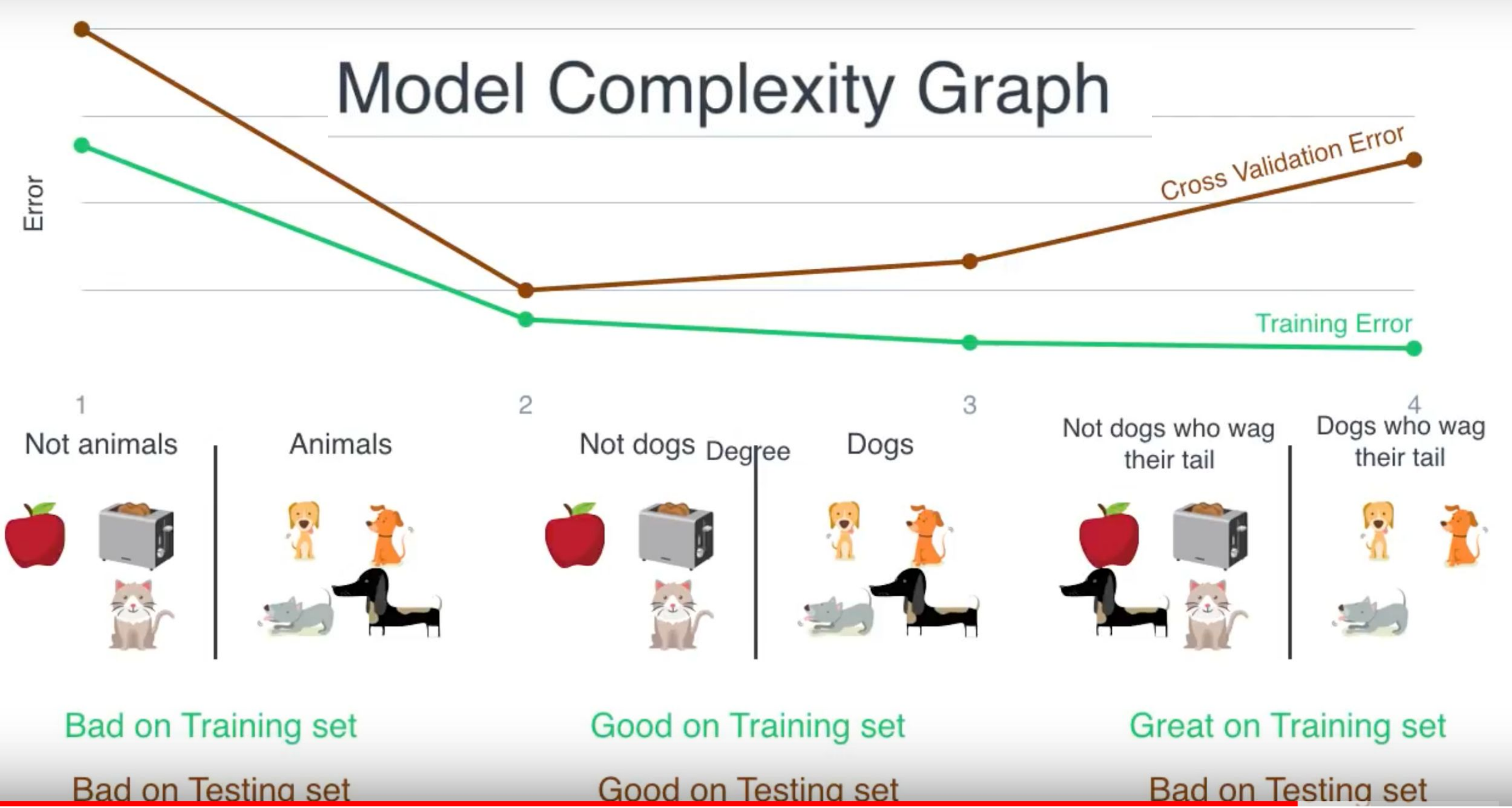
You should never use your testing data for training!!!

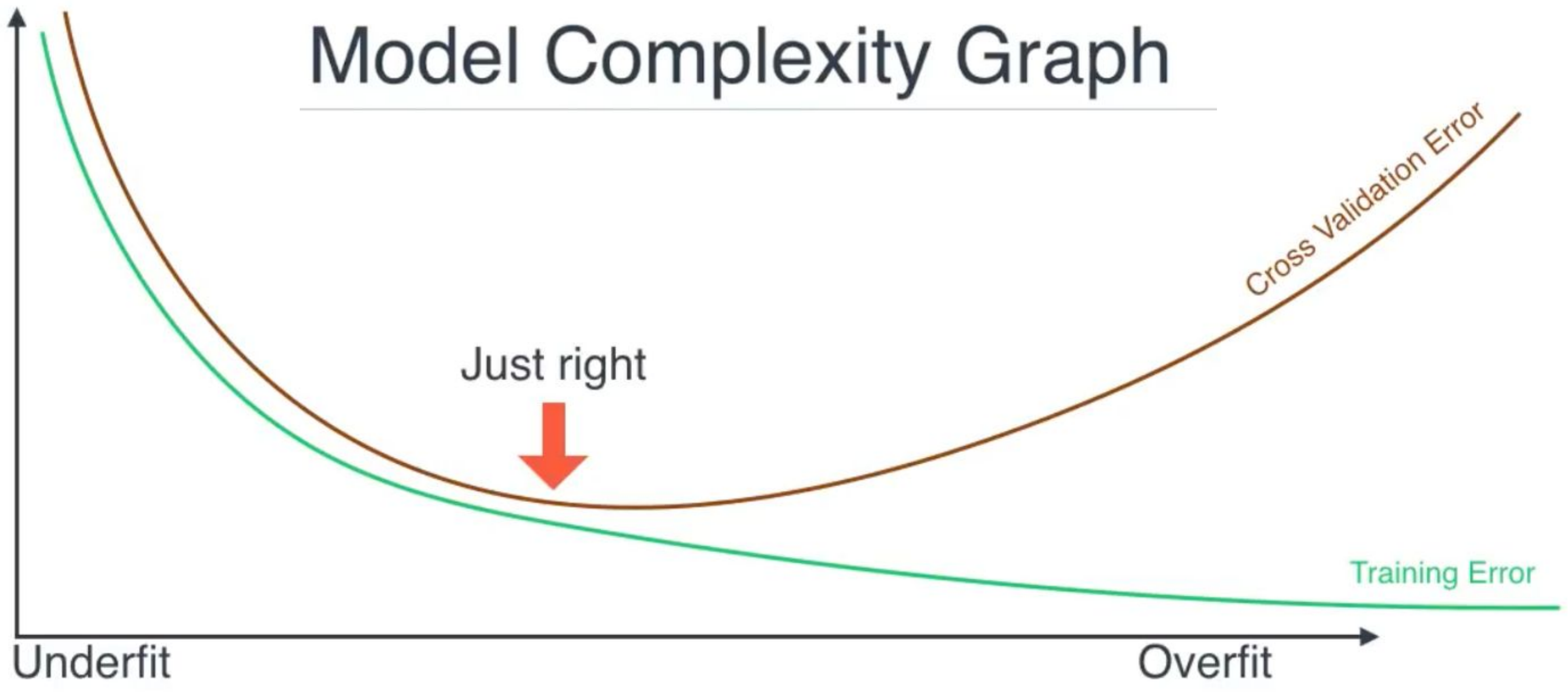


Training - Testing - Validation Datasets

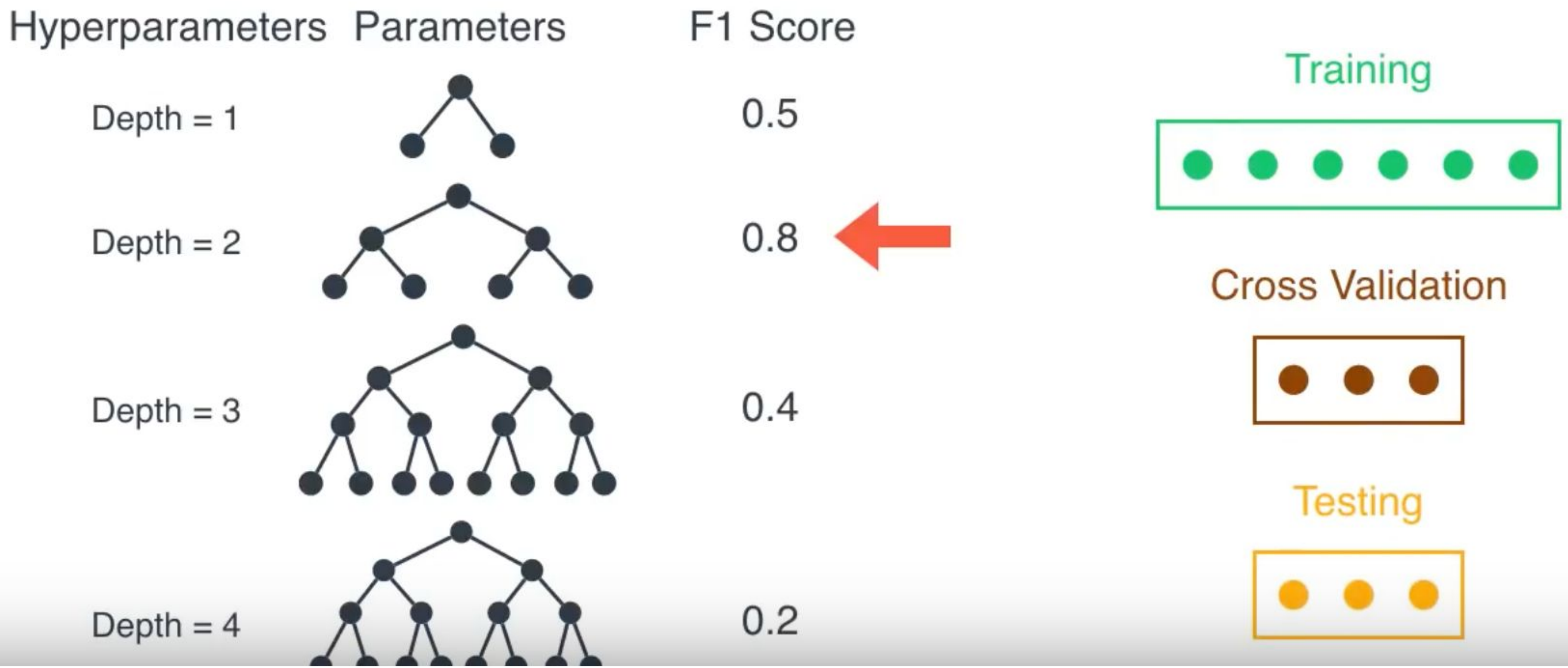
Solution: Cross Validation




Model Complexity Graph
Error
Cross Validation Error
Training Error
1
2
3
4
Degree
Not animals
Animals
Not dogs
Dogs
Not dogs who wag their tail
Dogs who wag their tail
Bad on Training set
Good on Training set
Great on Training set
Bad on Testing set
Good on Testing set
Bad on Testing set

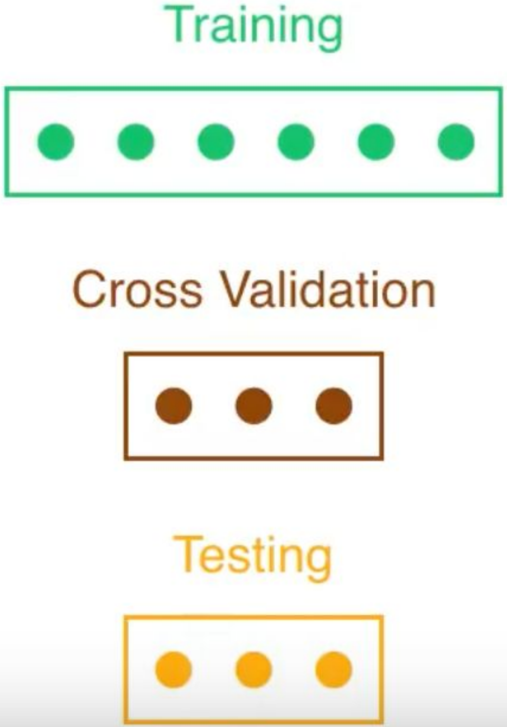

Model Complexity Graph
Cross Validation Error
Just right
Training Error
Underfit
Overfit

Training a Decision Tree



Training a Logistic Regression Model

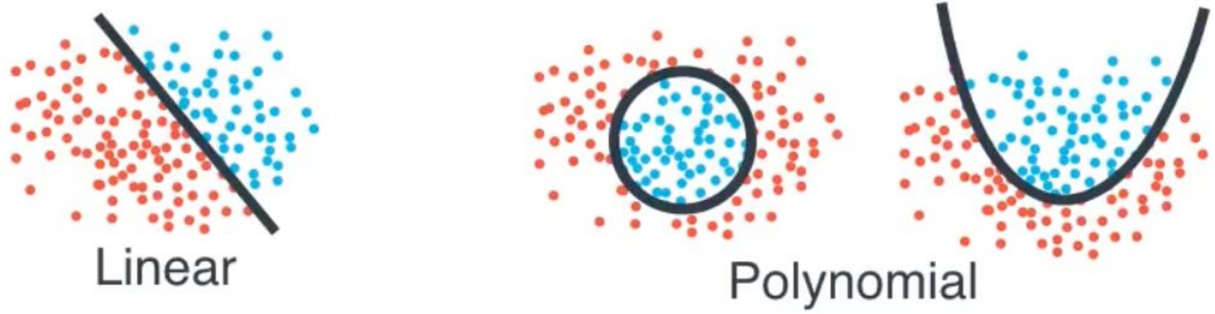
	Parameters	F1 Score
Degree = 1		0.5
Degree = 2		0.8
Degree = 3		0.4
Degree = 4		0.2



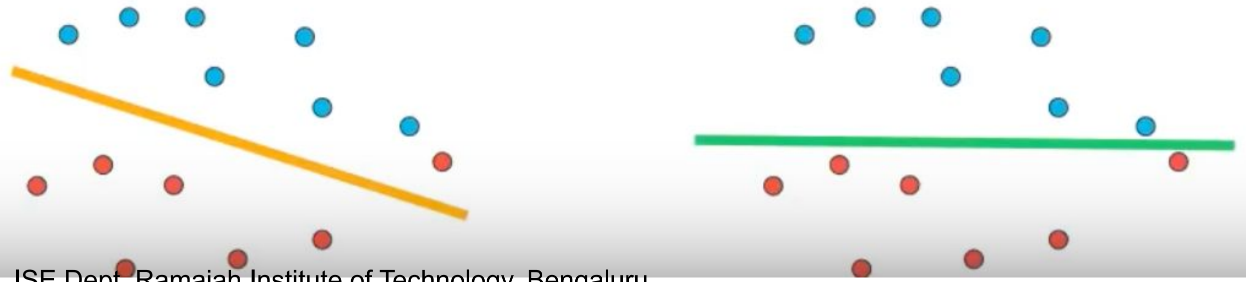
Training a Support Vector Machine

Hyperparameters

Kernel



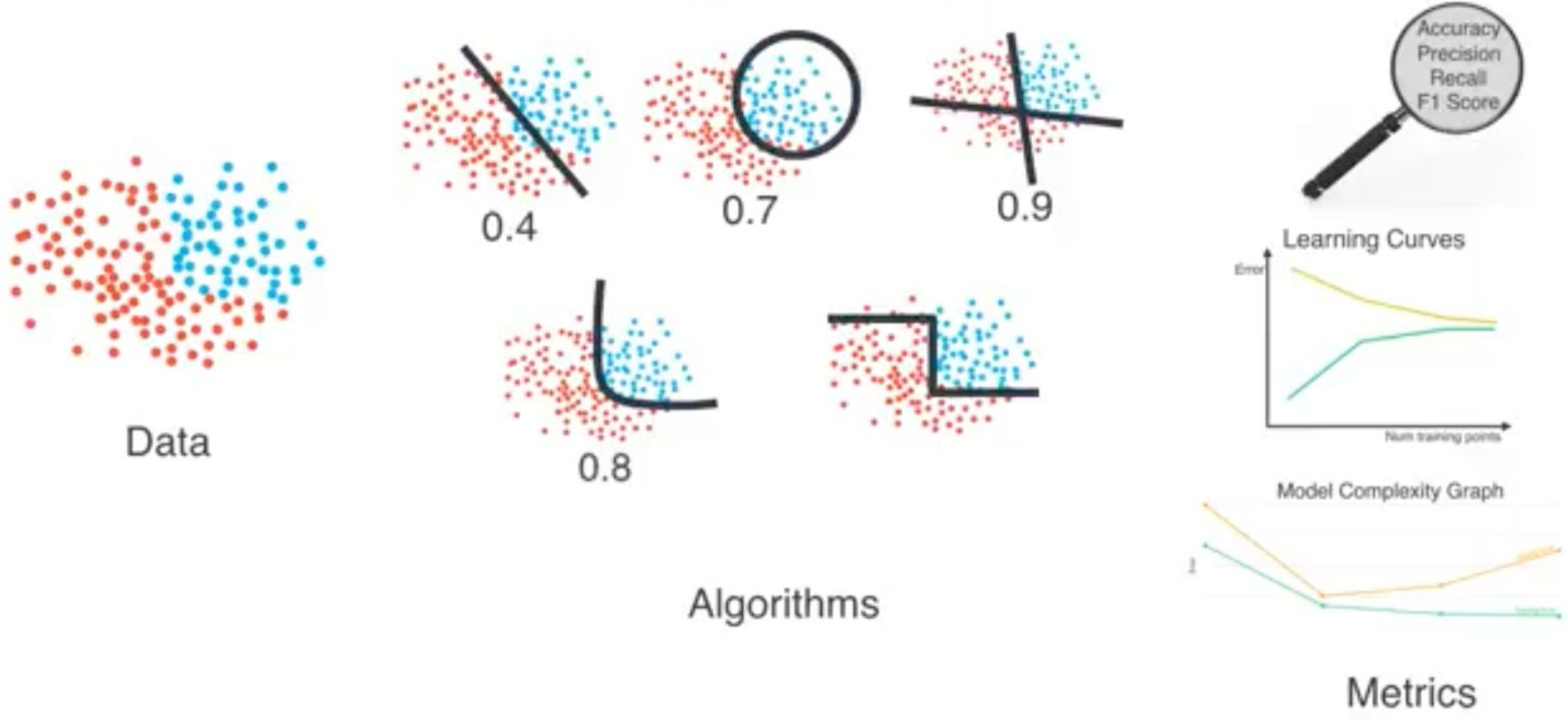
γ



Parameters and Hyperparameters

Algorithm	Parameters	Hyperparameters
Random Forest	Features Thresholds	Number of trees Depth
Logistic Regression	Coefficients of the polynomial	Degree of the polynomial
Support Vector Machines	Coefficients	Kernel Gamma C

How to use machine learning



Source:

1. **Bias Variance Dichotomy:** https://www.youtube.com/watch?v=DtZ3NaPNBNE&feature=youtu.be
2. **Train-Test-Validate Dataset:** https://www.youtube.com/watch?v=e2vurxnd124&feature=youtu.be
3. **(Song)Stat Quest - Bias Variance :** https://youtu.be/EuBBz3bI-aA