MIDS W207
Applied Machine Learning

Spring 2022

Week 2
Live Session Slides

Fundamentals

Supervised Learning

- Makes machine Learn explicitly
- Data with clearly defined output is given
- Direct feedback is given
- Predicts outcome/future
- Resolves classification and regression problems

Unsupervised Learning

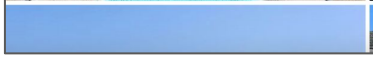
- Machine understands the data (Identifies patterns/structures)
- Evaluation is qualitative or indirect
- Does not predict/find anything specific

Reinforcement Learning

- An approach to AI
- Reward based learning
- Learning form +ve & +ve reinforcement
- Machine Learns how to act in a certain environment
- To maximize rewards

Prediction

Prediction



$740,000 4 bd | 3 ba | 2,028 sqft

30198 La Primavera St, Temecula, CA 92592

Est. payment: $3,399/mo Get pre-qualified

Message

Overview | Facts and features | Home value | Price and tax hist

Facts and features

Type: Single Family Residence

Year built: 1974

Heating: Central

Cooling: Central Air

Parking: 2 Attached Garage spaces

HOA: $90 monthly

Lot: 0.40 Acres

Price/sqft: $365

Interior details

Bedrooms and bathrooms

Bedrooms: 4
Bathrooms: 3
Full bathrooms: 2
1/2 bathrooms: 1
Main level bathrooms: 1

Flooring

Flooring: Carpet, Laminate, Tile

Appliances

Appliances included: Built-In Range, Gas Cooktop, Disposal, Refrigerator
Laundry features: Inside, Laundry Room

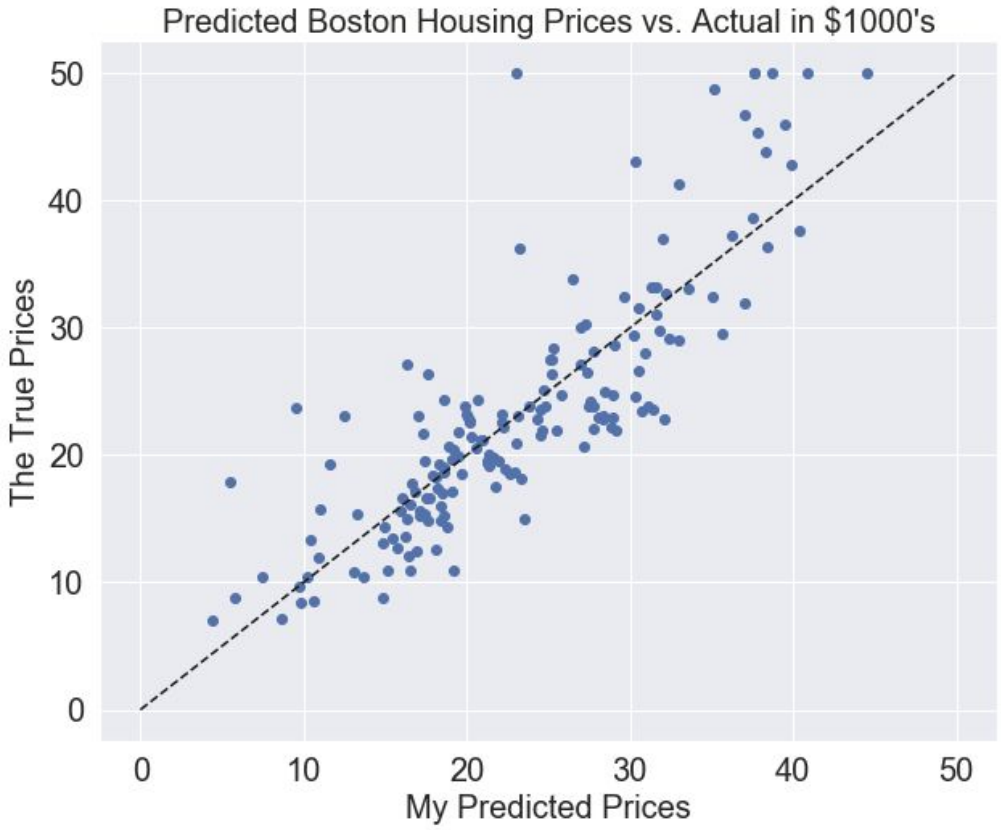
Interior Features

Interior features: Ceiling Fan(s), Granite Counters, All Bedrooms Up

Prediction: Boston Real Estate

	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	PTRATIO	B	LSTAT	MEDV
0	0.00632	18.0	2.31	0.0	0.538	6.575	65.2	4.0900	1.0	296.0	15.3	396.90	4.98	24.0
1	0.02731	0.0	7.07	0.0	0.469	6.421	78.9	4.9671	2.0	242.0	17.8	396.90	9.14	21.6
2	0.02729	0.0	7.07	0.0	0.469	7.185	61.1	4.9671	2.0	242.0	17.8	392.83	4.03	34.7
3	0.03237	0.0	2.18	0.0	0.458	6.998	45.8	6.0622	3.0	222.0	18.7	394.63	2.94	33.4
4	0.06905	0.0	2.18	0.0	0.458	7.147	54.2	6.0622	3.0	222.0	18.7	396.90	5.33	36.2
5	0.02985	0.0	2.18	0.0	0.458	6.430	58.7	6.0622	3.0	222.0	18.7	394.12	5.21	28.7

Prediction: Boston Real Estate



Prediction: Boston Real Estate

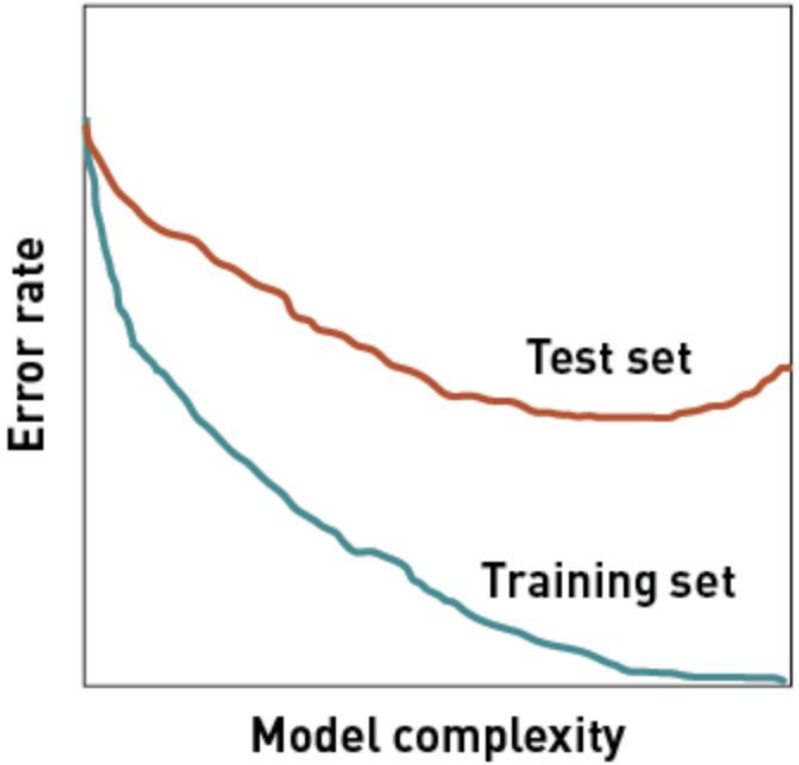
Coefficient of Determination or R-squared

$$R^2 = 1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \bar{y})^2}$$

Root Mean Squared Error

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2}$$

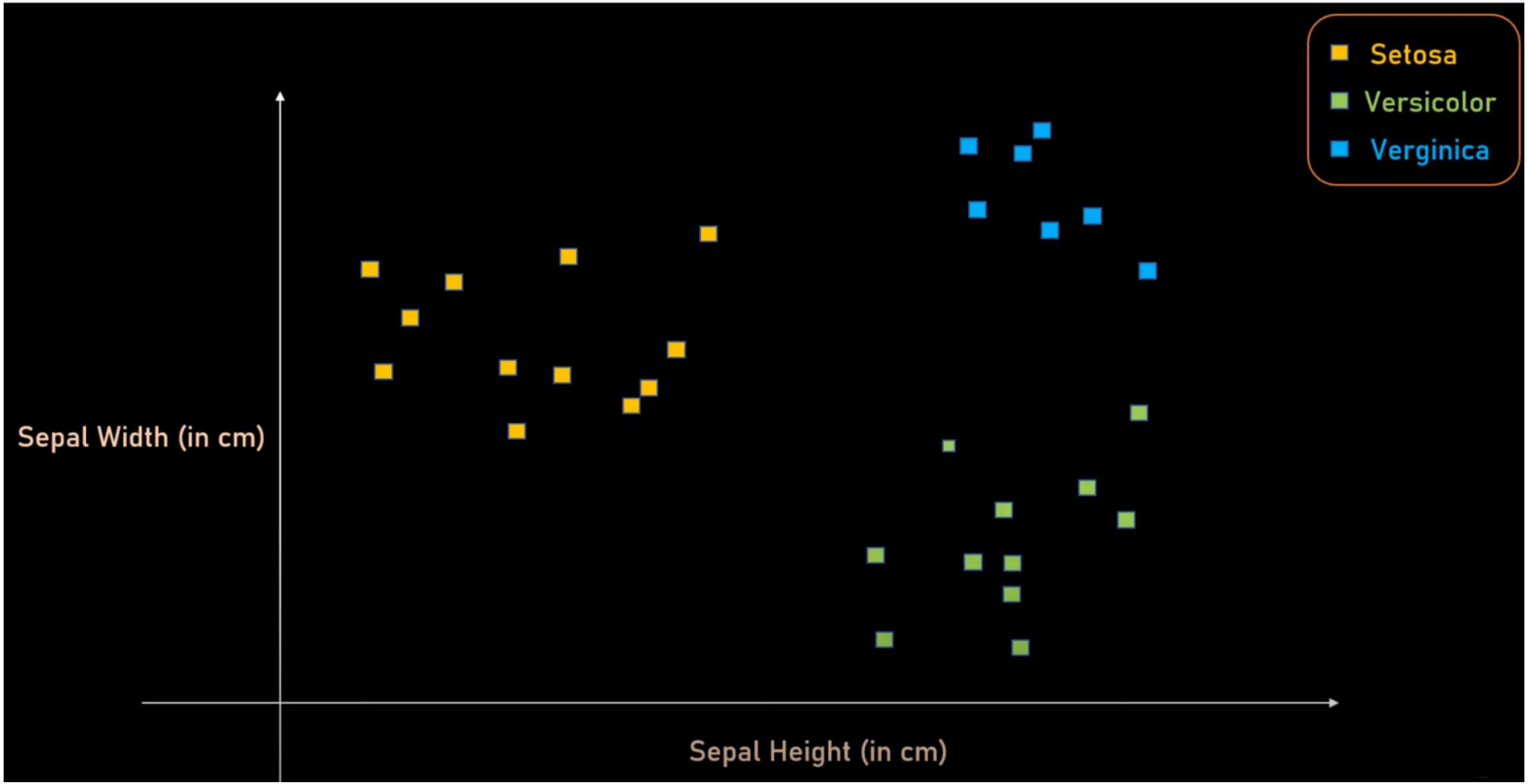
Training and Test Data



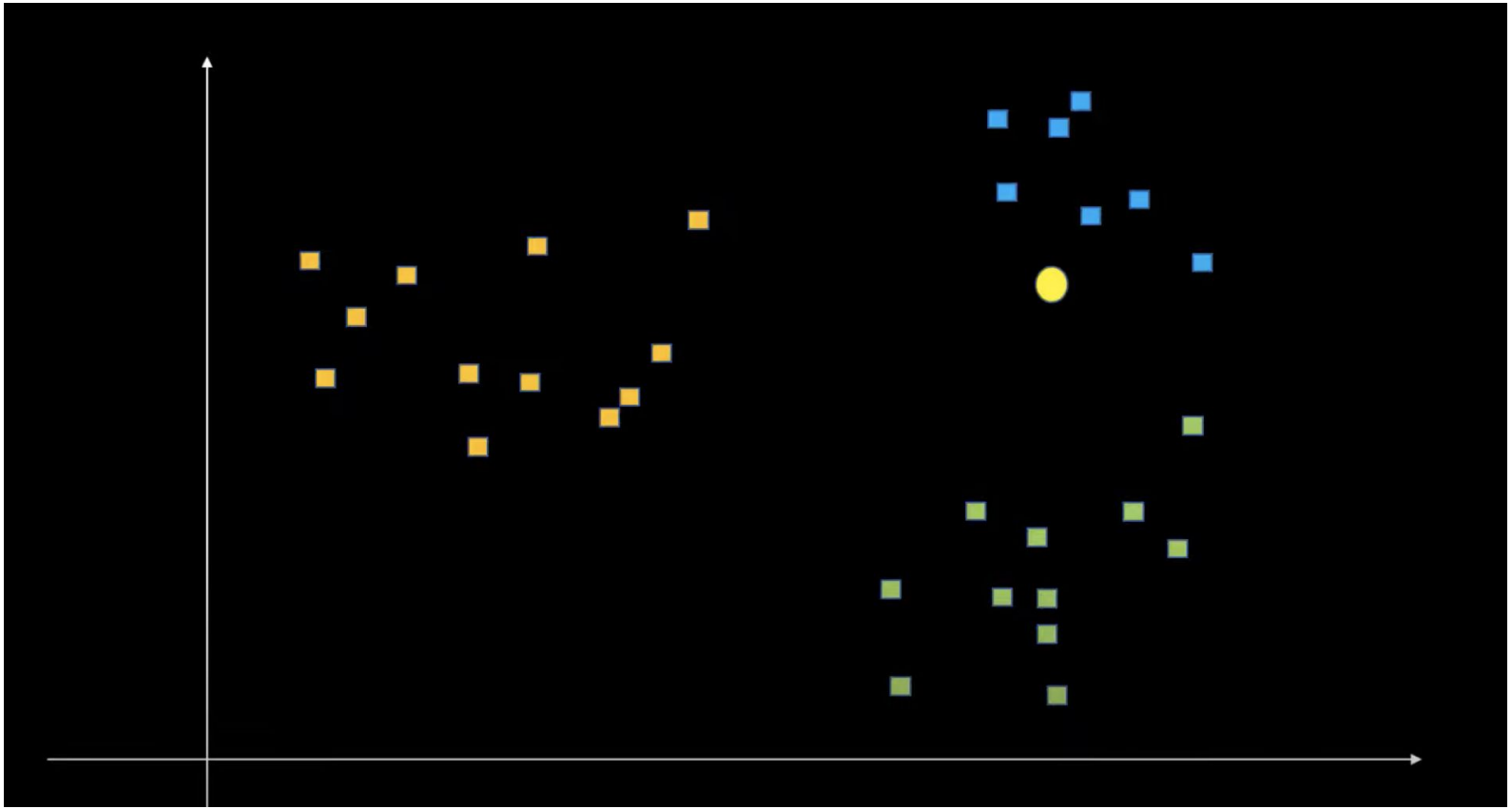
K- Nearest Neighbors Algorithm

Setosa **Versicolor** **Virginica**

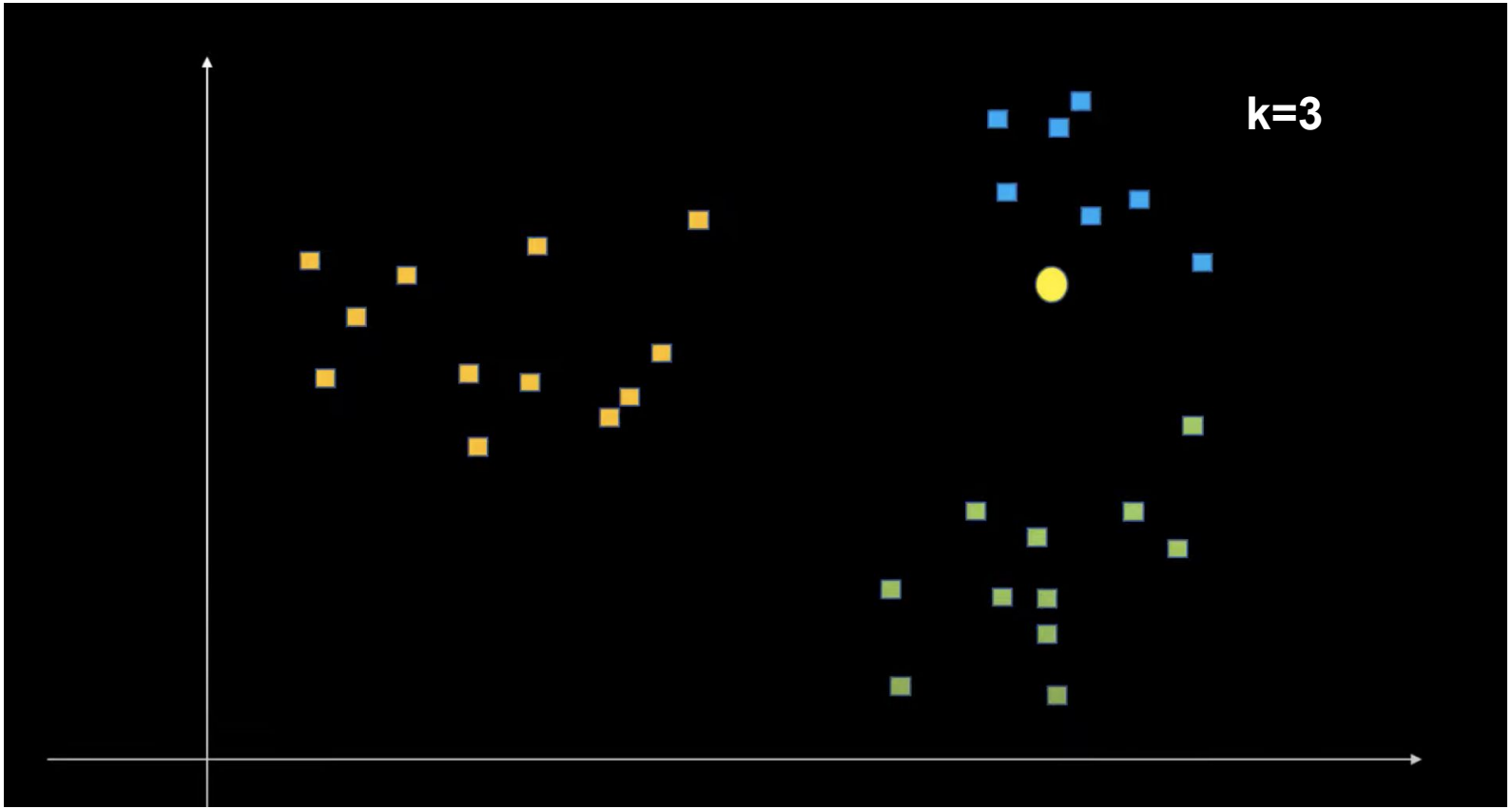
K- Nearest Neighbors Algorithm



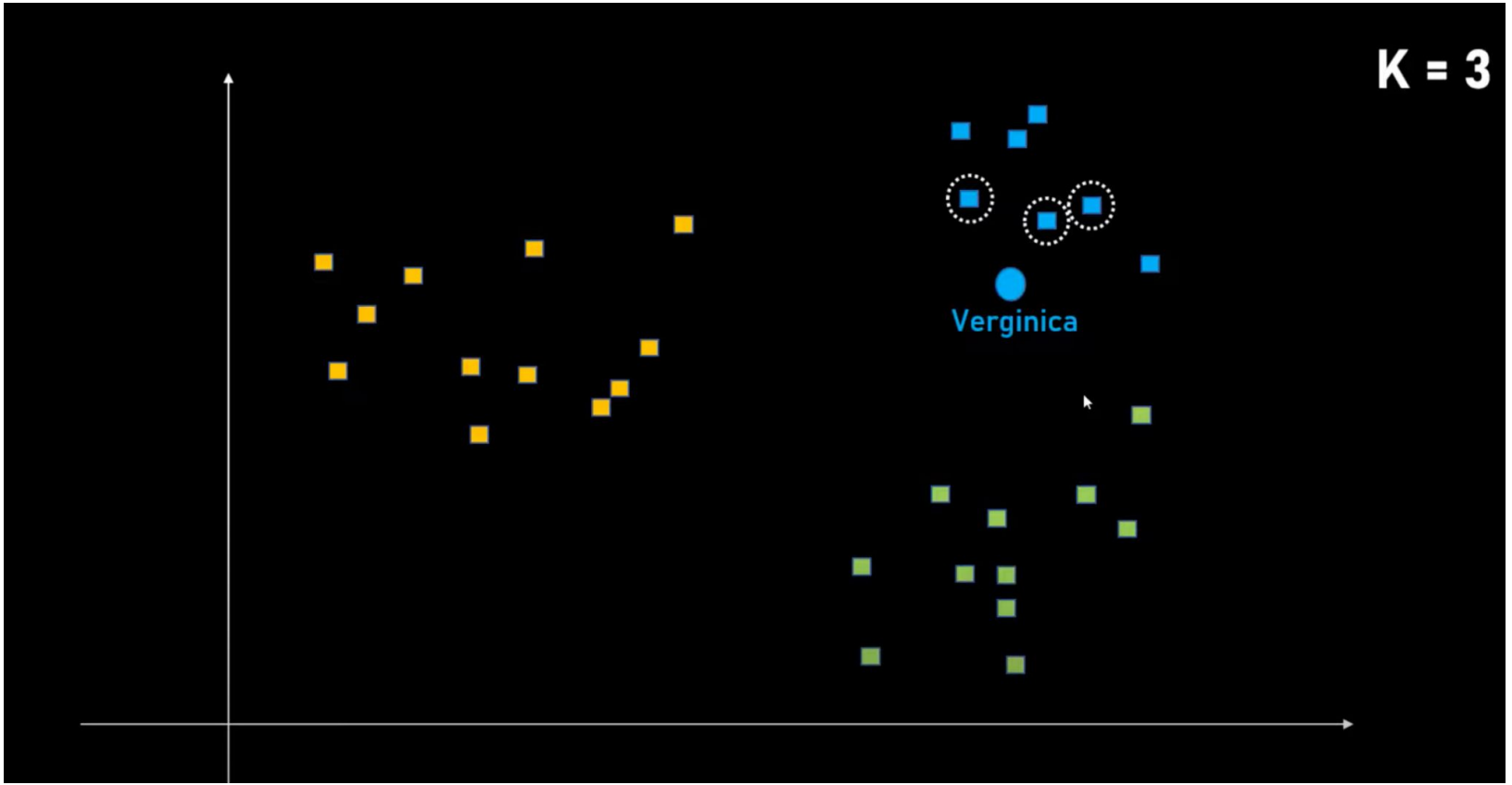
K- Nearest Neighbors Algorithm



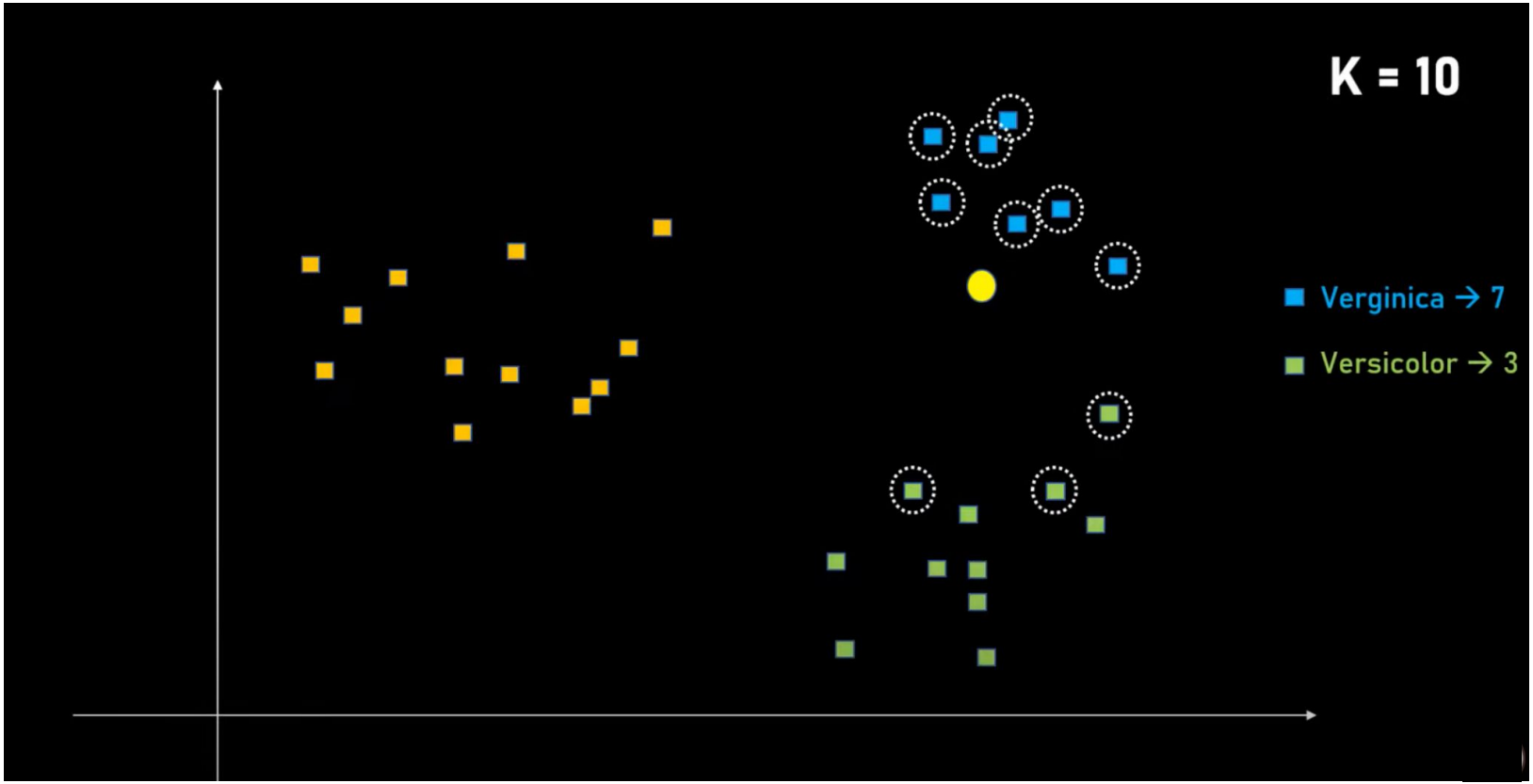
K- Nearest Neighbors Algorithm



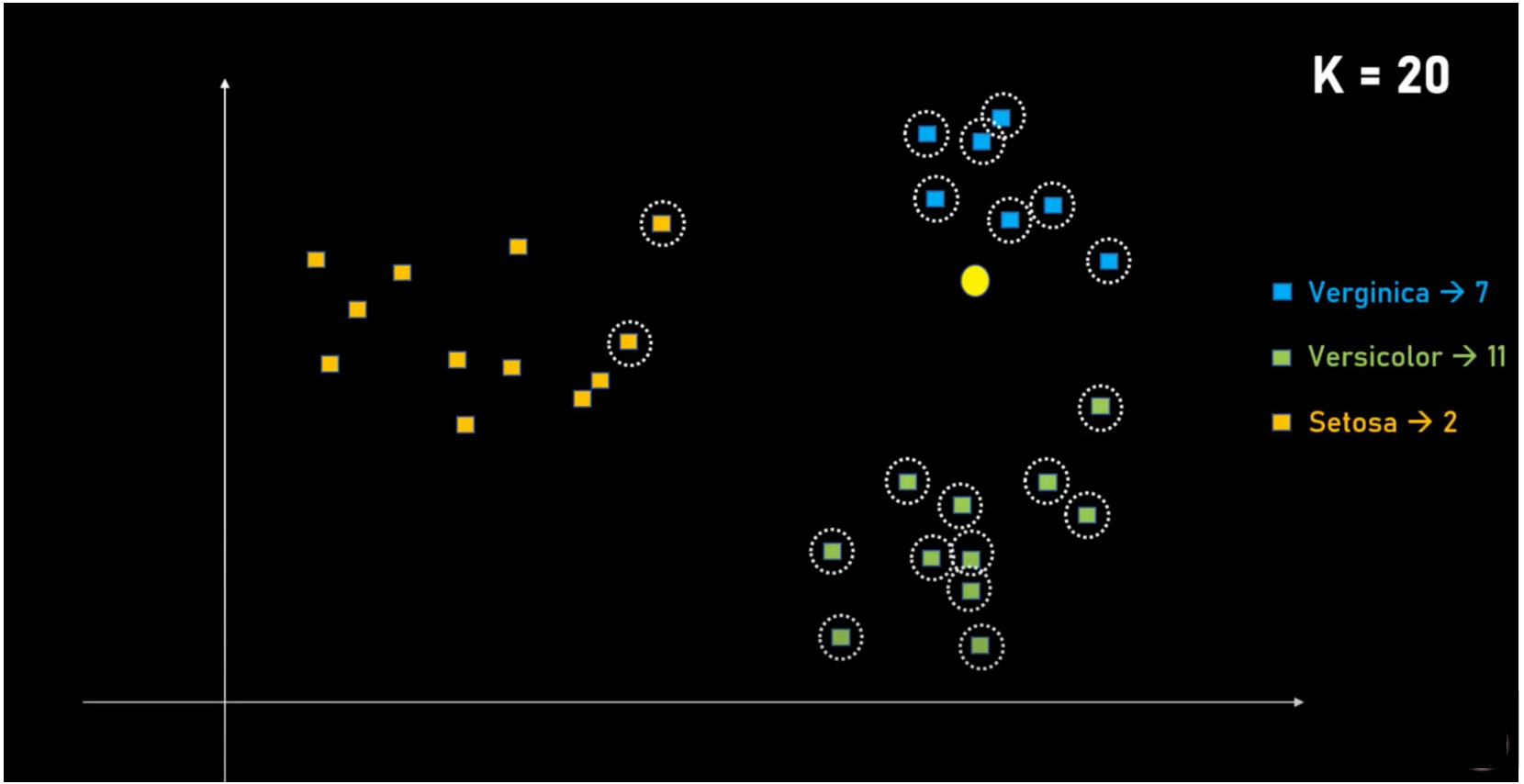
K- Nearest Neighbors Algorithm



K- Nearest Neighbors Algorithm

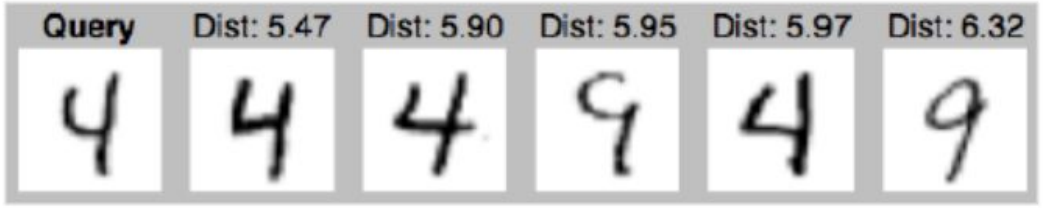


K- Nearest Neighbors Algorithm

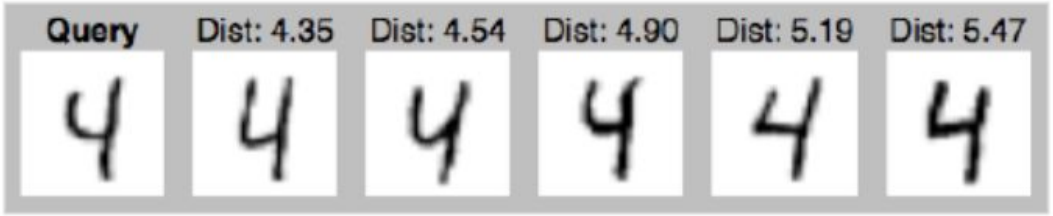


Digit Classification

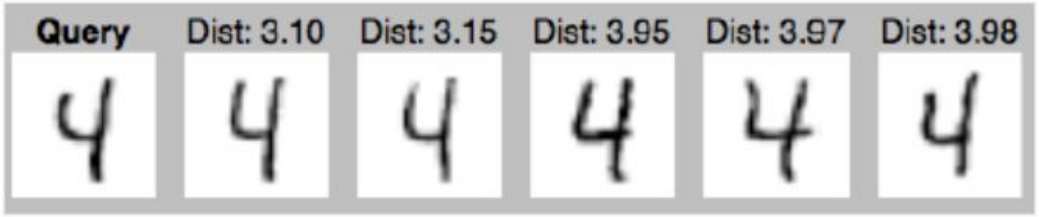
200 examples



1,000 examples



10,000 examples



Digit Classification

- Test set size = 10,000 digits

k = 1; Euclidean (L2) distance

Training	Error %	Time (secs)
100	30.0	0.38
1,000	12.1	2.34
10,000	5.3	28.7
60,000	2.7	2202
Deskewing	2.3	
Blurring	1.8	
Pixel shifting	1.2	

Code Review