Applying Clustering to Image Data

Janani Ravi
CO-FOUNDER, LOONYCORN

www.loonycorn.com

Overview

Working with image data

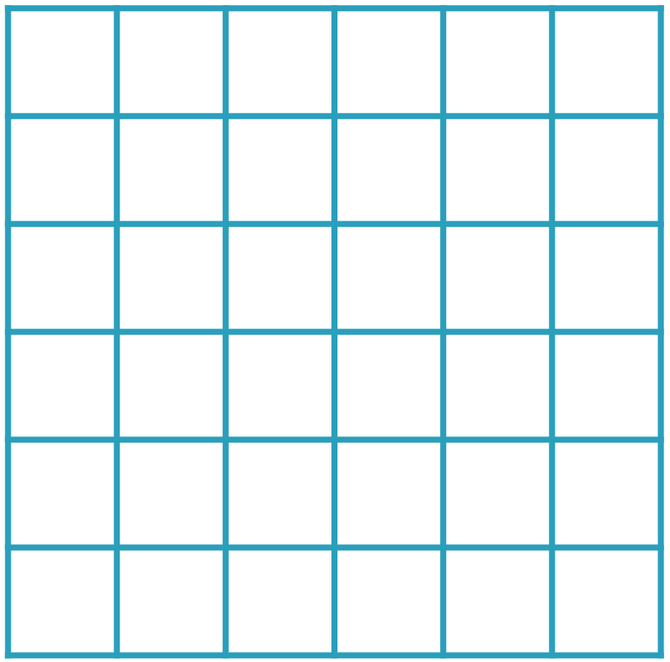
Converting images to feature vectors

Partitioning image data using clusters

Classifying images using clusters

Images as Matrices

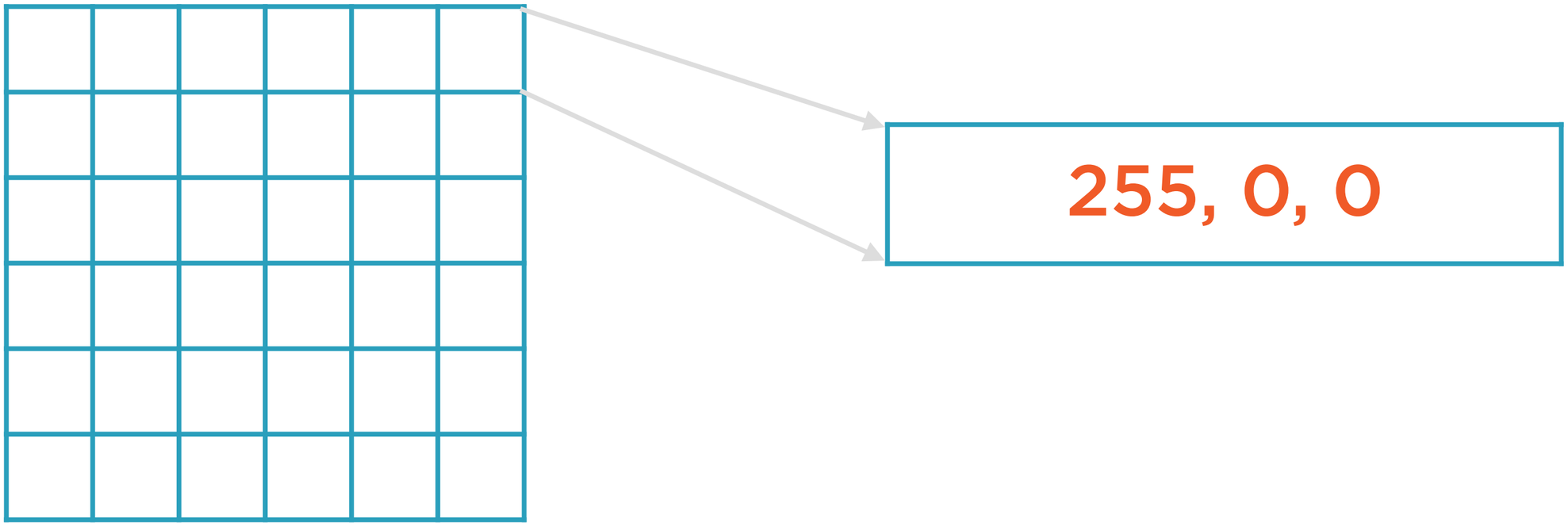
RGB Images



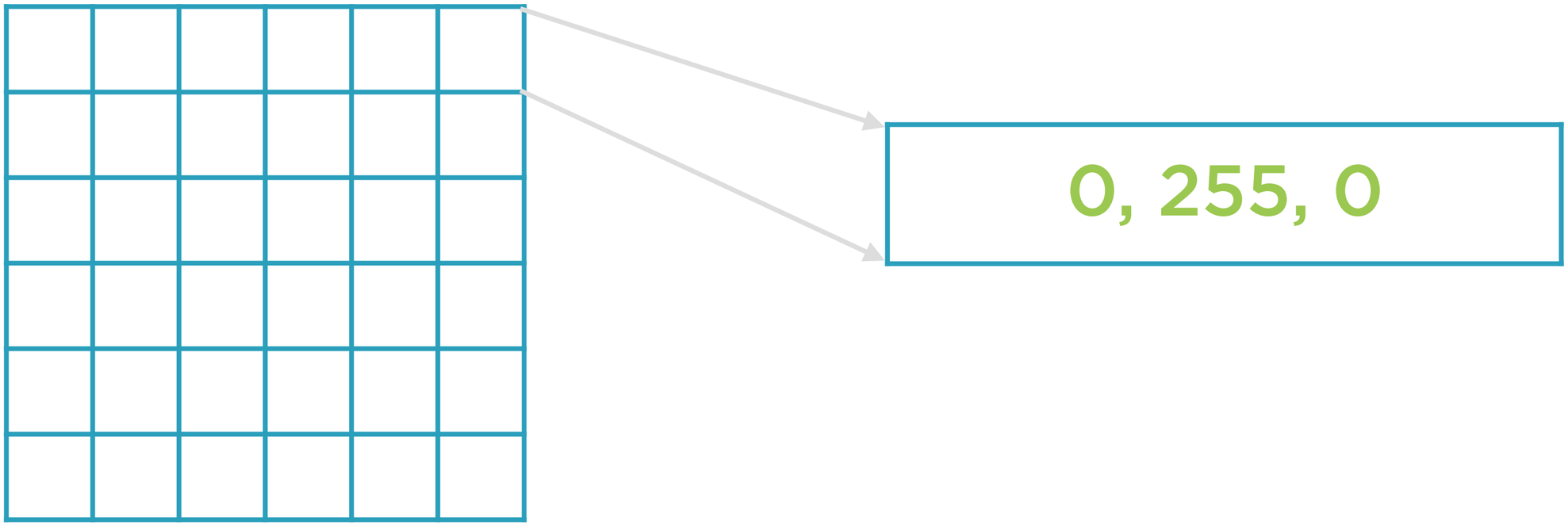
RGB values are for color images

R, G, B: 0-255

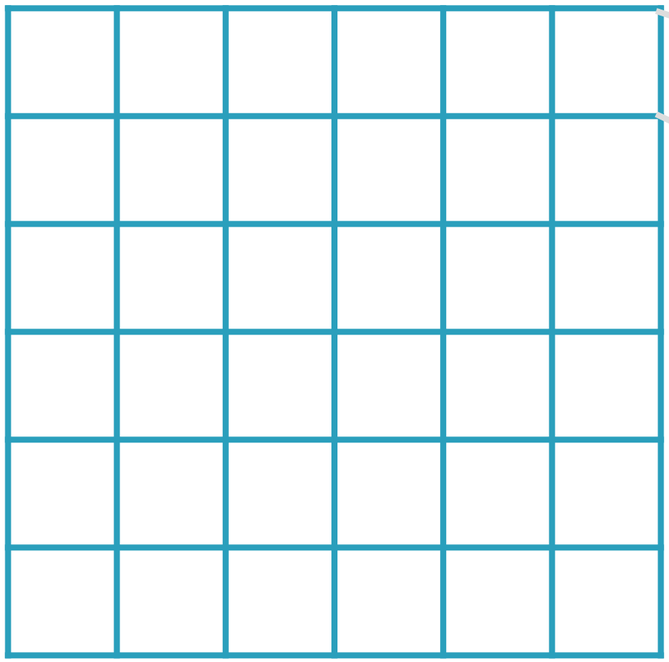
RGB Images



RGB Images



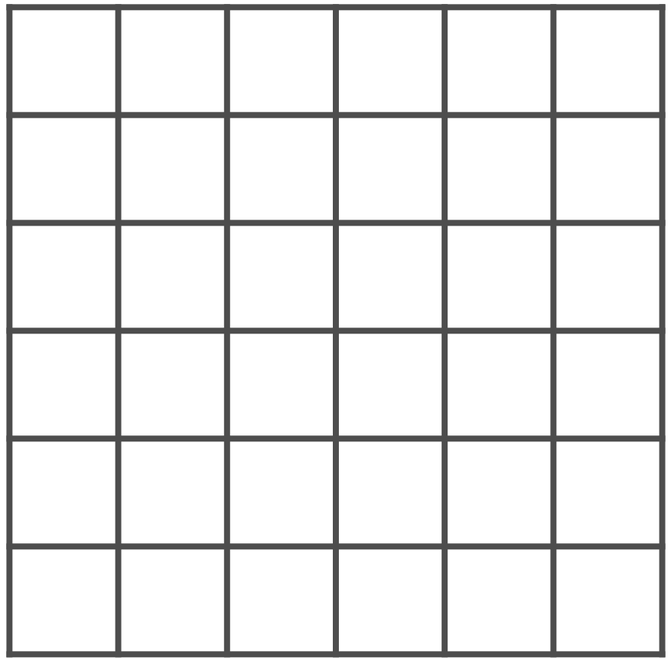
RGB Images



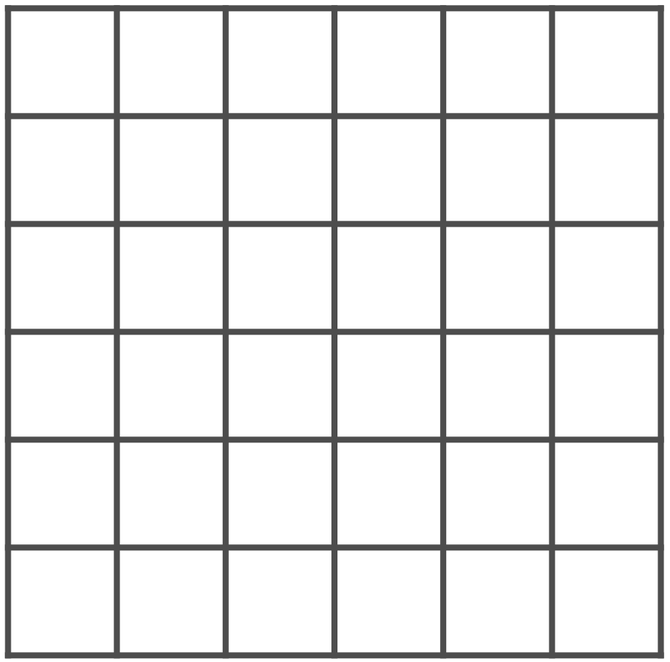
0, 0, 255

3 values to represent color, 3 channels

Grayscale Images



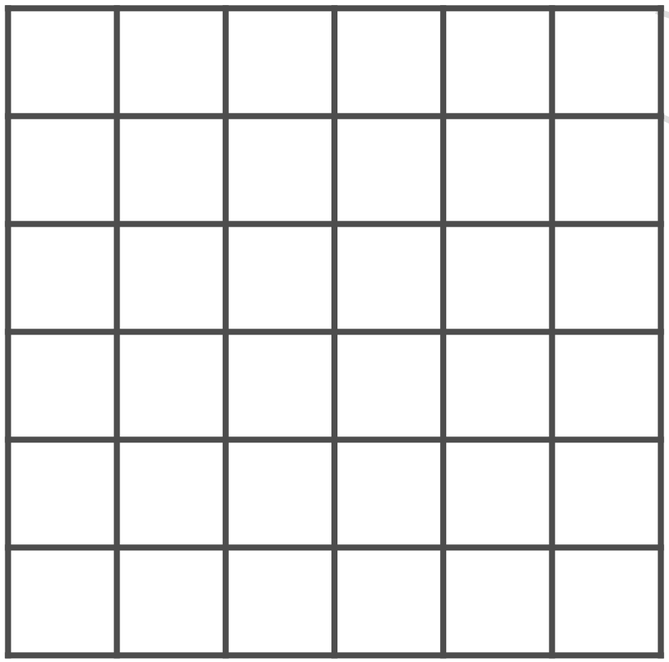
Grayscale Images



Each pixel represents only intensity information

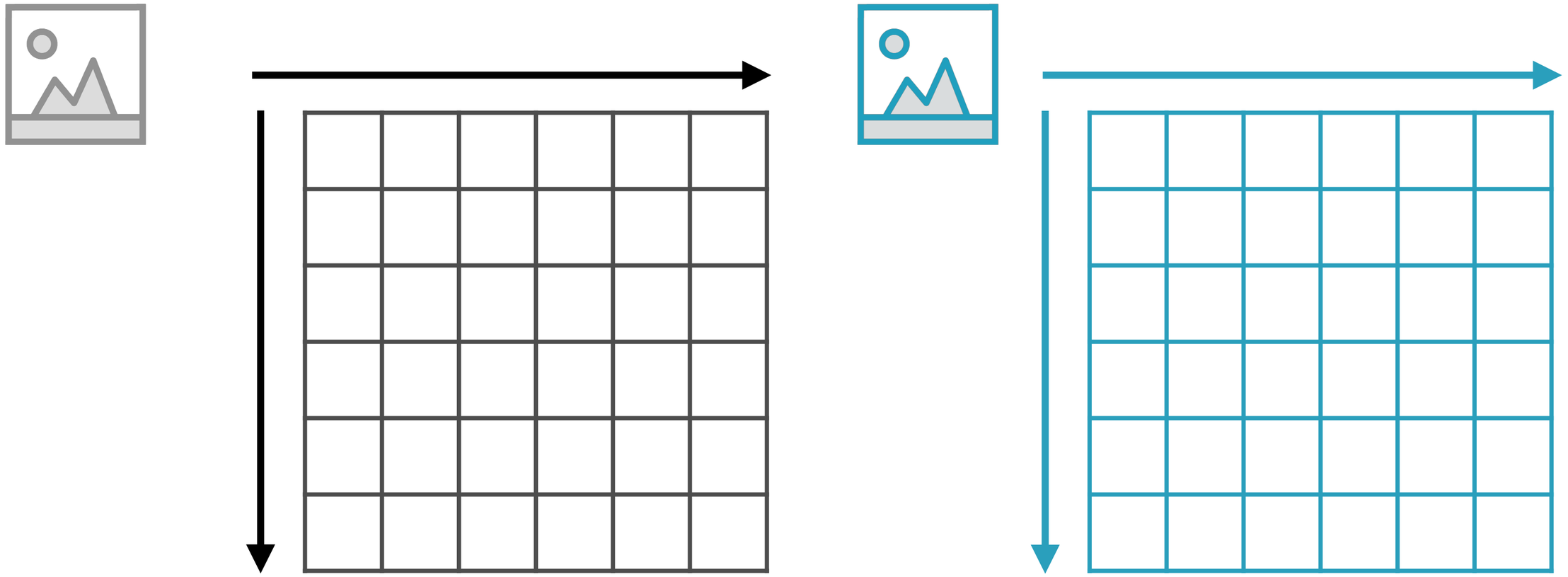
0.0 - 1.0

Grayscale Images



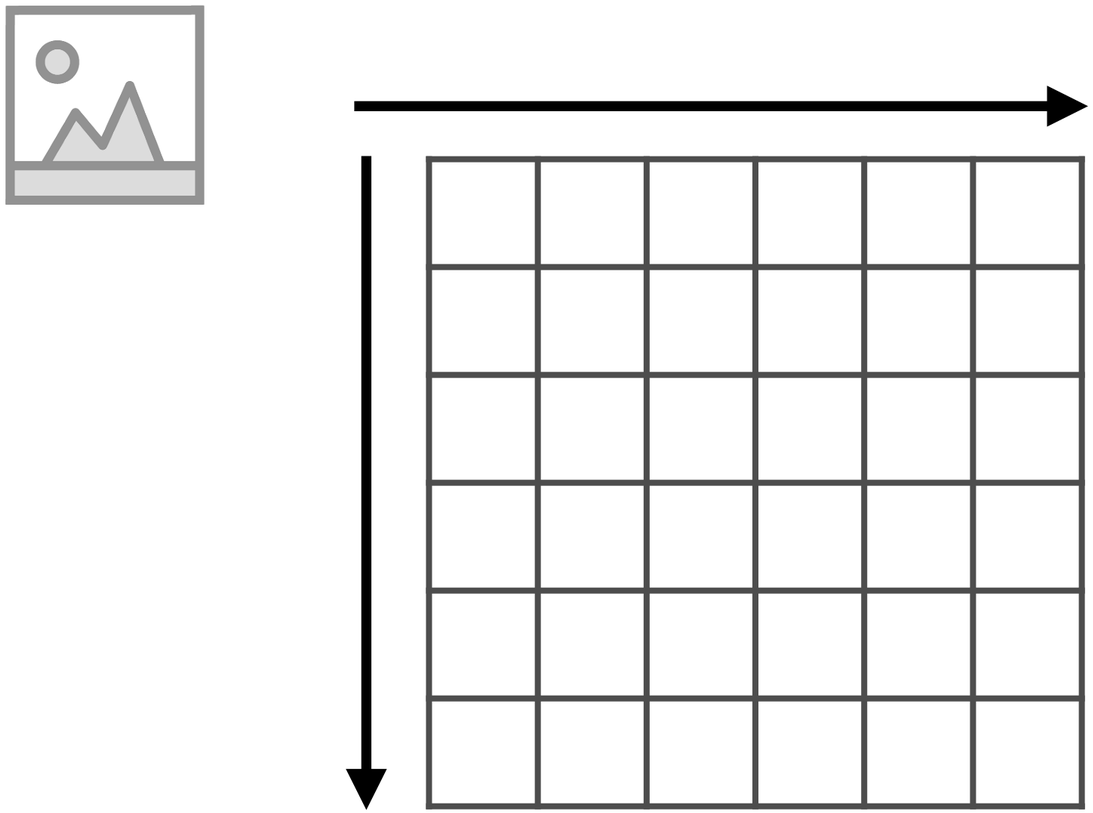
1 value to represent intensity, 1 channel

Images as Matrices

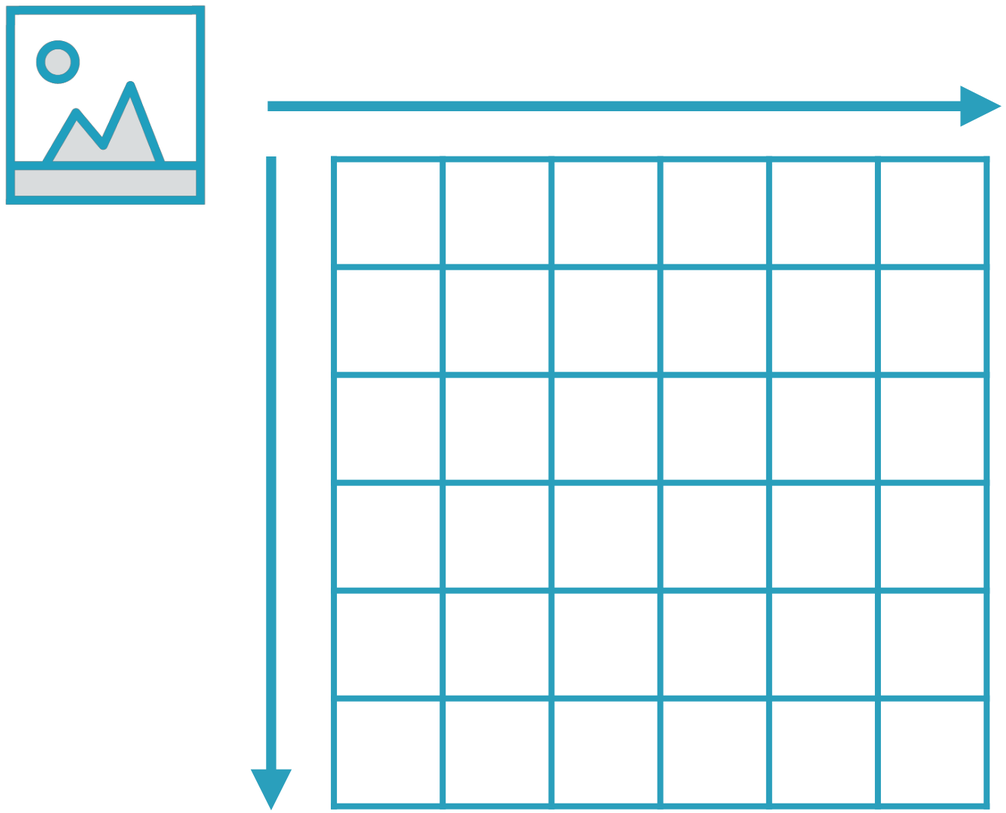


Images can be represented by a 3-D matrix

Images as Matrices



(6, 6, 1)



(6, 6, 3)

List of Images

ML frameworks (e.g. TensorFlow) usually deal with a list of images in one 4-D Tensor

List of Images

The images should all be the same size

List of Images

(10, 6, 6, 3)

The number of channels

List of Images

(10, 6, 6, 3)

The height and width of each image in the list

List of Images

(10, 6, 6, 3)

The number of images

Demo

Clustering images in the MNIST handwritten dataset

Summary

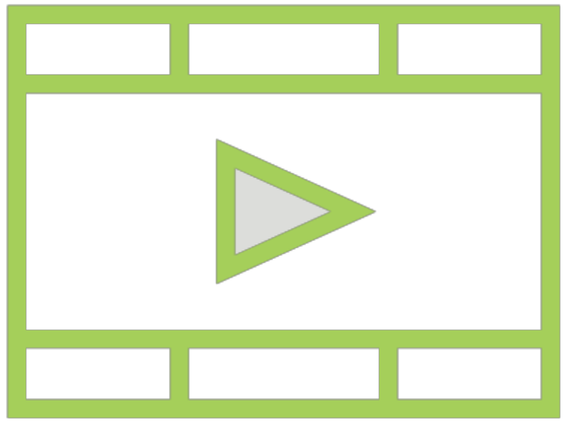
Working with image data

Converting images to feature vectors

Partitioning image data using clusters

Classifying images using clusters

Related Courses



Employing Ensemble Methods with scikit-learn

Foundations of PyTorch

Reducing Dimensions in Data with scikit-learn