Implementing Text and Image Classification Using Neural Networks in scikit-learn

Janani Ravi
CO-FOUNDER, LOONYCORN

www.loonycorn.com

Overview

Representing text as features

Term Frequency (TF) and Inverse Document Frequency (IDF)

Representing images as features

Single channel and multi-channel images

Text and image classification

Encoding Text Data in Numeric Form

d = "This is not the worst restaurant in the metropolis, not by a long way"
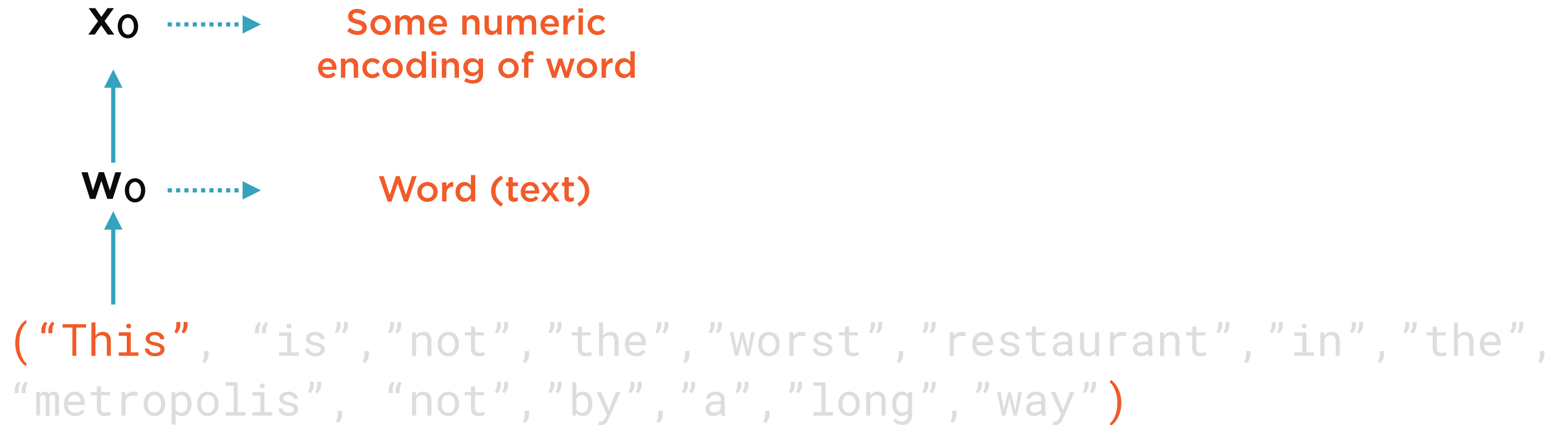
Document as Word Sequence

Model a document as an ordered sequence of words

```
d = "This is not the worst restaurant in the metropolis,
not by a long way"

("This", "is","not","the","worst","restaurant","in","the",
"metropolis", "not","by","a","long","way")
```

Document as Word Sequence

Tokenize document into individual words

X_0 ┈┈┈▶ **Some numeric encoding of word**

W_0 ┈┈┈▶ **Word (text)**

("This", "is","not","the","worst","restaurant","in","the",
"metropolis", "not","by","a","long","way")

Represent Each Word as a Number

X_1 ⟶ Some numeric encoding of word

W_1 ⟶ Word (text)

("This", "is","not","the","worst","restaurant","in","the", "metropolis", "not","by","a","long","way")

Represent Each Word as a Number

X_n ┈┈➤ Some numeric encoding of word

W_n ┈┈➤ Word (text)

("This", "is","not","the","worst","restaurant","in","the", "metropolis", "not","by","a","long","way")
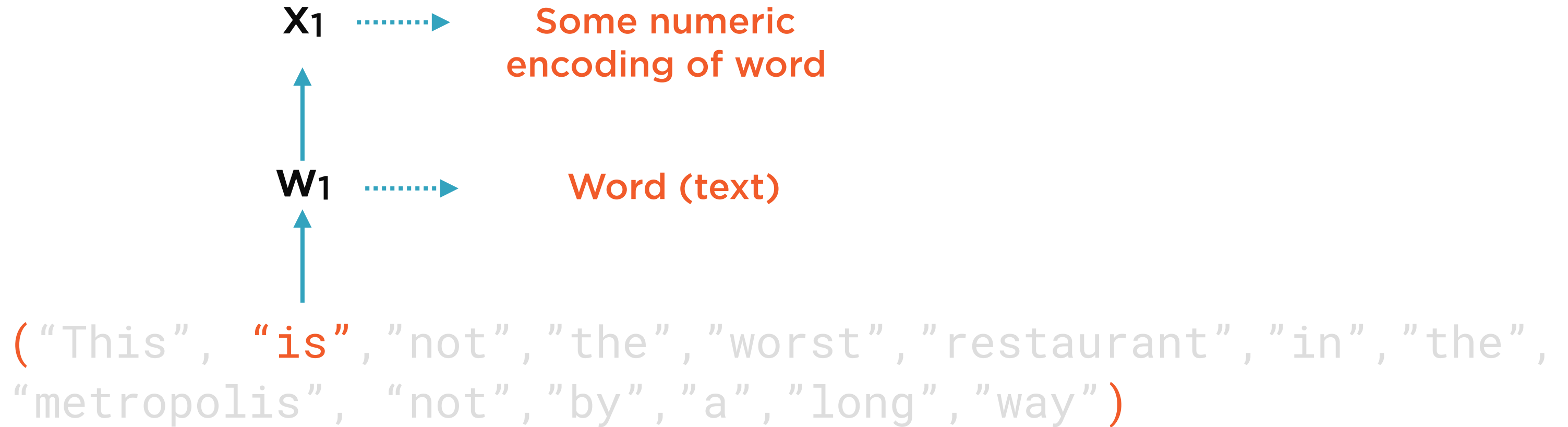
Represent Each Word as a Number

$$d = [x_0, x_1, \ldots x_n]$$

Document as Tensor

Represent each word as numeric data, aggregate into tensor

Numeric Representations of Text

One-hot

Frequency-based

Prediction-based

Numeric Representations of Text

One-hot

Frequency-based

Prediction-based

Represent each word in text by its presence or absence

Numeric Representations of Text

One-hot

Frequency-based

Prediction-based

Frequency-based Embeddings

| Count | TF-IDF | Co-occurrence |

Frequency-based Embeddings

Count

TF-IDF

Co-occurrence

Capture how often a word occurs in a document i.e. the **counts** or the **frequency**

Frequency-based Embeddings

Count

TF-IDF

Co-occurrence

Captures how often a word occurs in a **document** as well as the **entire corpus**

Tf-Idf

Frequently in a single document

Might be important
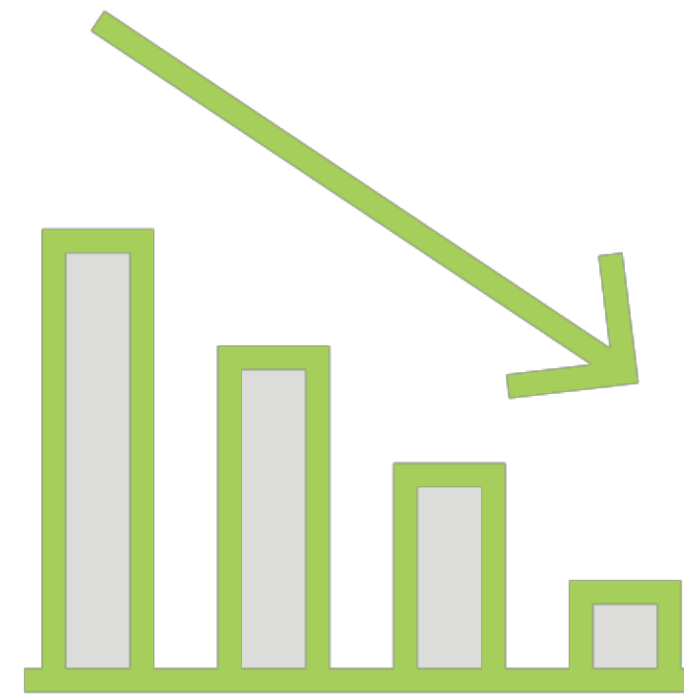
Frequently in the corpus

Probably a common word like "a", "an", "the"

Frequency-based Embeddings

Count	TF-IDF	Co-occurrence

Similar words will occur together and will have similar context

Context Window

A window centered around a word, which includes a certain number of neighboring words

Co-occurrence

The number of times two words $w1$ and $w2$ have occurred together in a context window

Word Embeddings

One-hot

Frequency-based

Prediction-based

Predictions-based embeddings

Numerical representations of text which capture meanings and semantic relationships, generated using ML models

Magic

Word embeddings capture meaning

"Queen" ~ "King" == "Woman" ~ "Man"

"Paris" ~ "France" == "London" ~ "England"

Dramatic dimensionality reduction

Demo

Performing classification on text data using neural networks in scikit-learn

Encoding Image Data in Numeric Form

Image Recognition
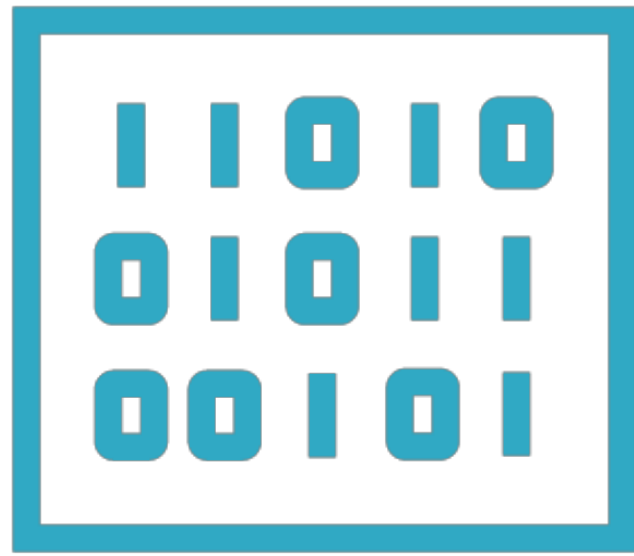
Images represented as pixels

Identify edges, colors, shapes

A photo of a horse

Encoding Image Data in Numeric Form

Images as Matrices

RGB Images
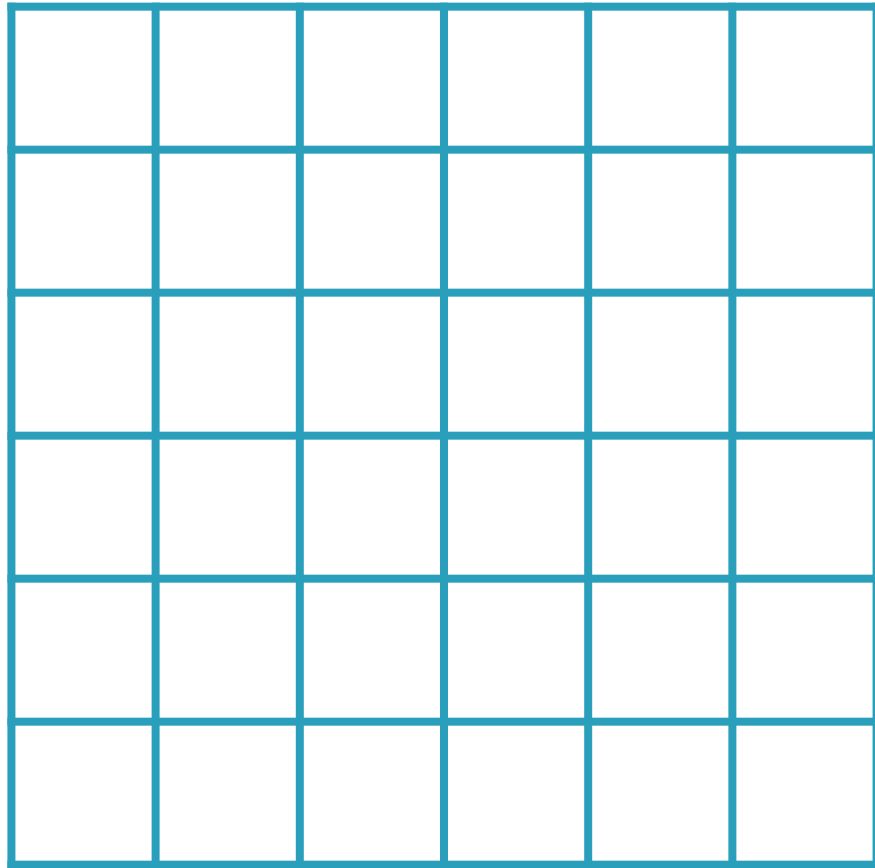
RGB values are for color images

R, G, B: 0-255

RGB Images
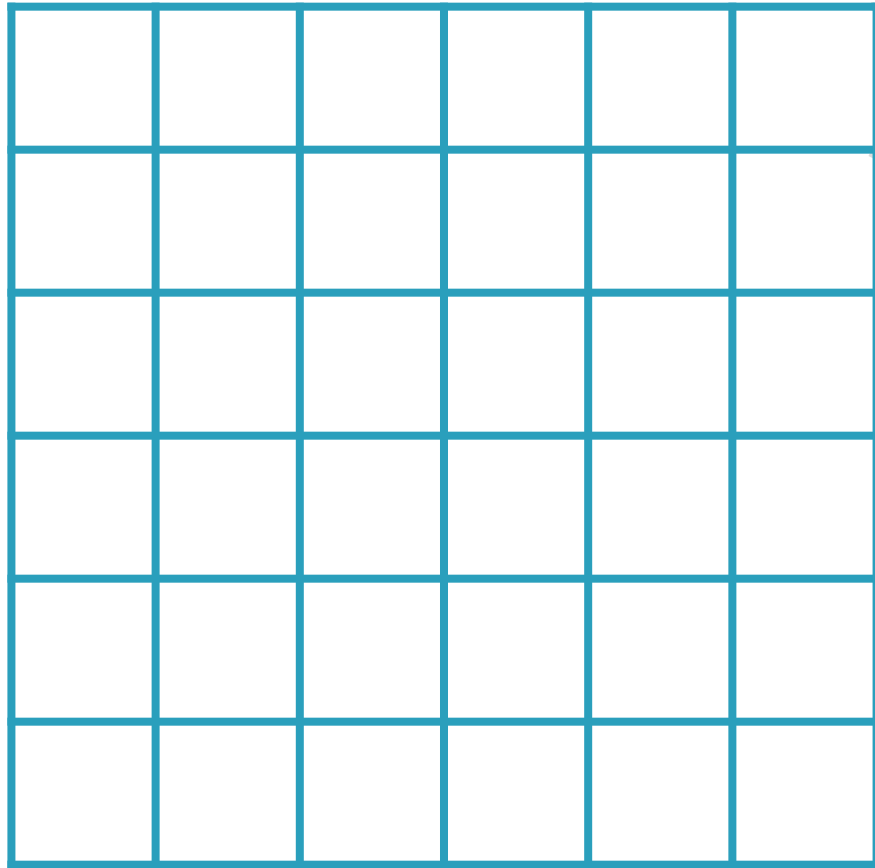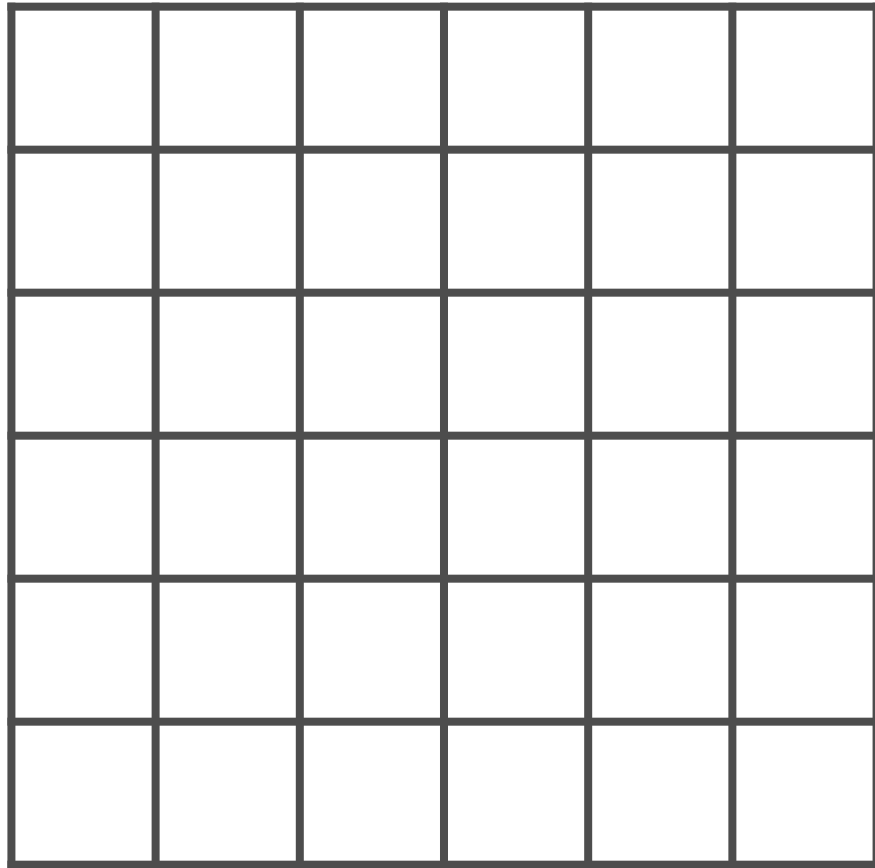
255, 0, 0

RGB Images
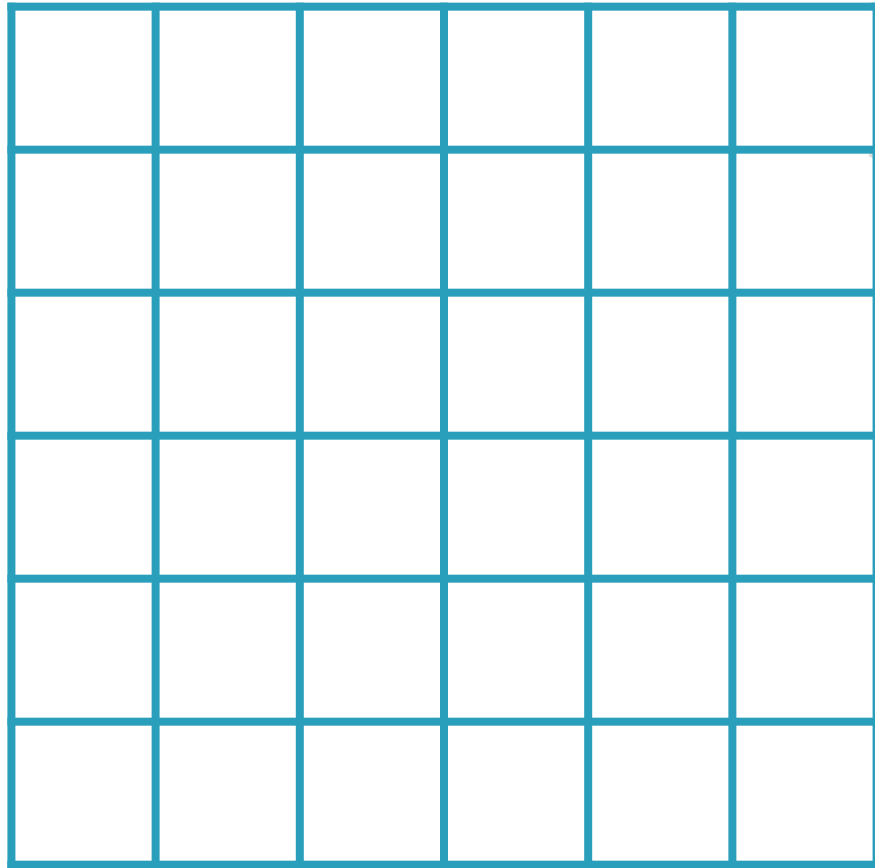
0, 255, 0

RGB Images
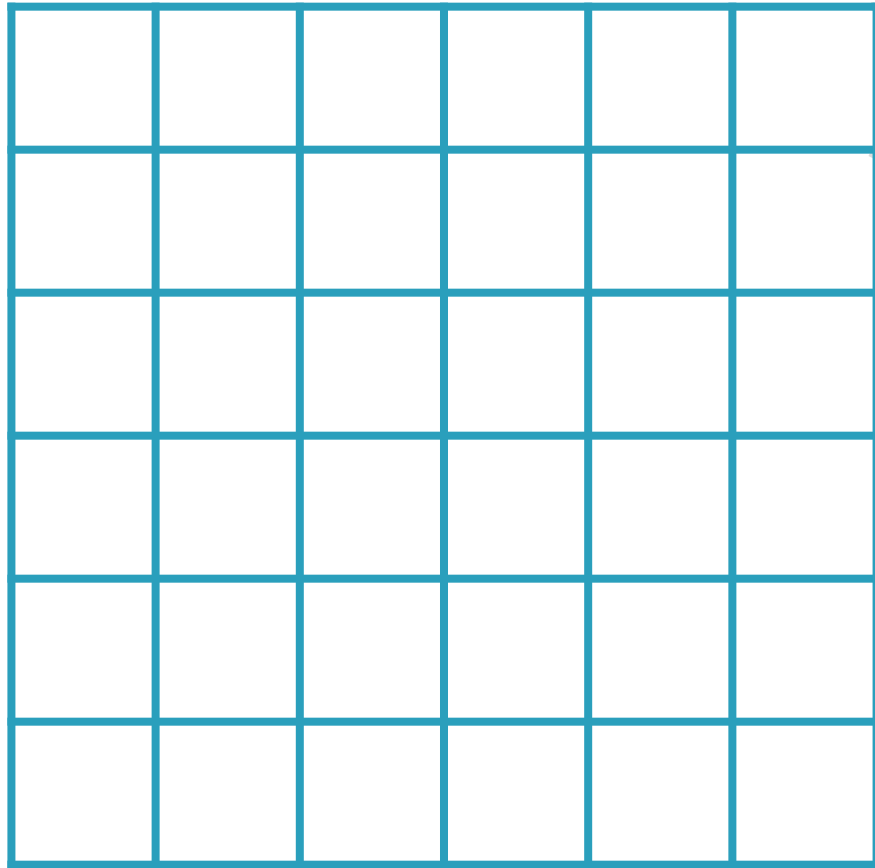
0, 0, 255

3 values to represent color, **3** channels

Grayscale Images
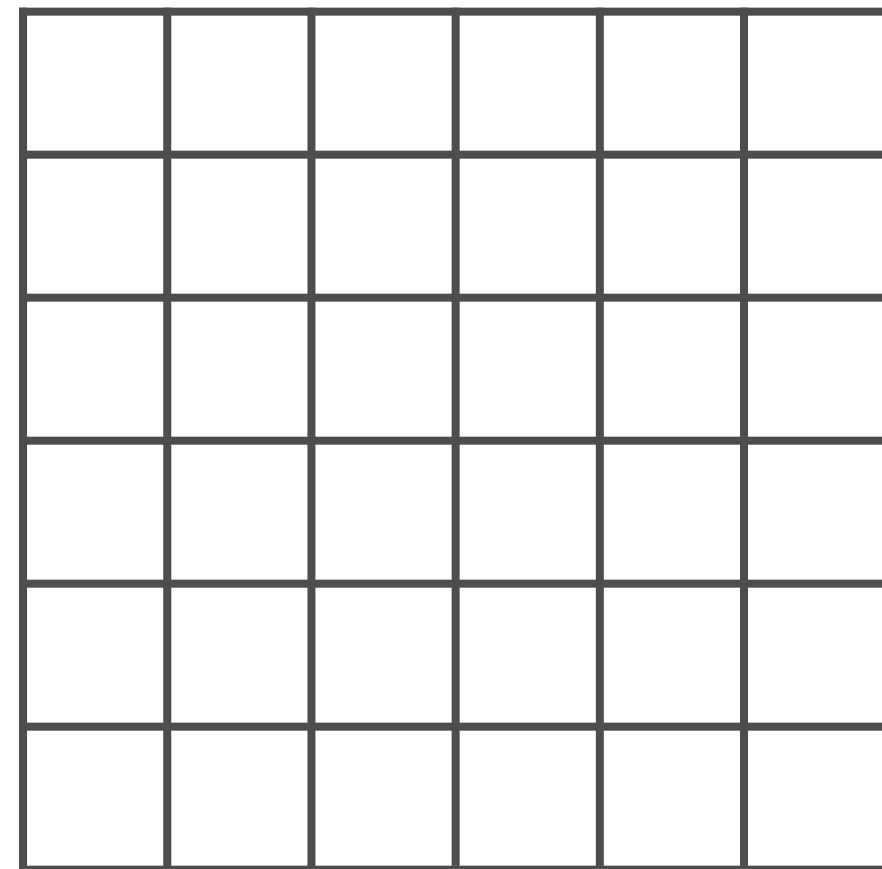
Grayscale Images

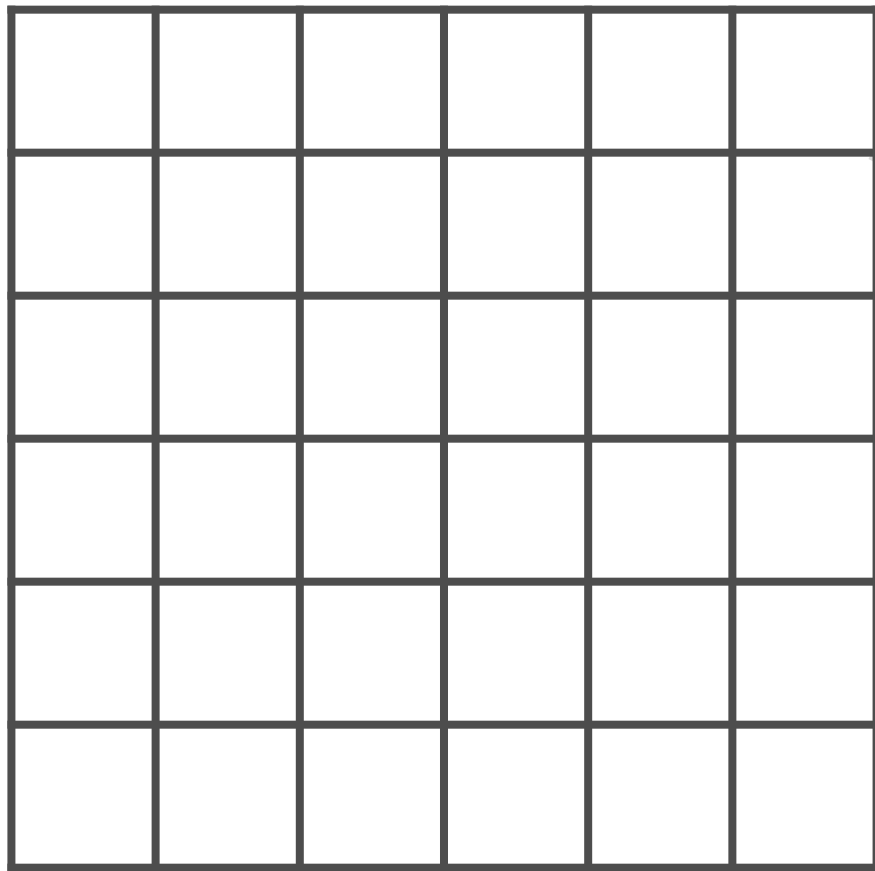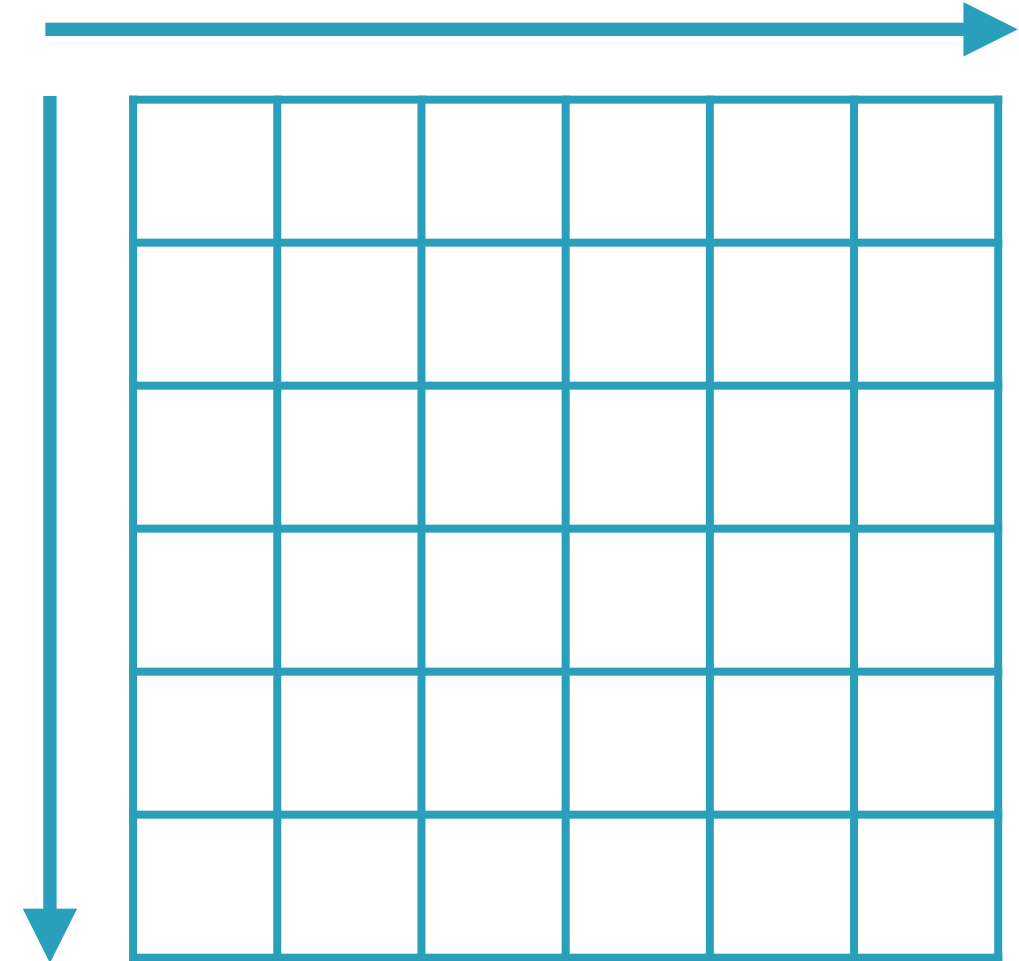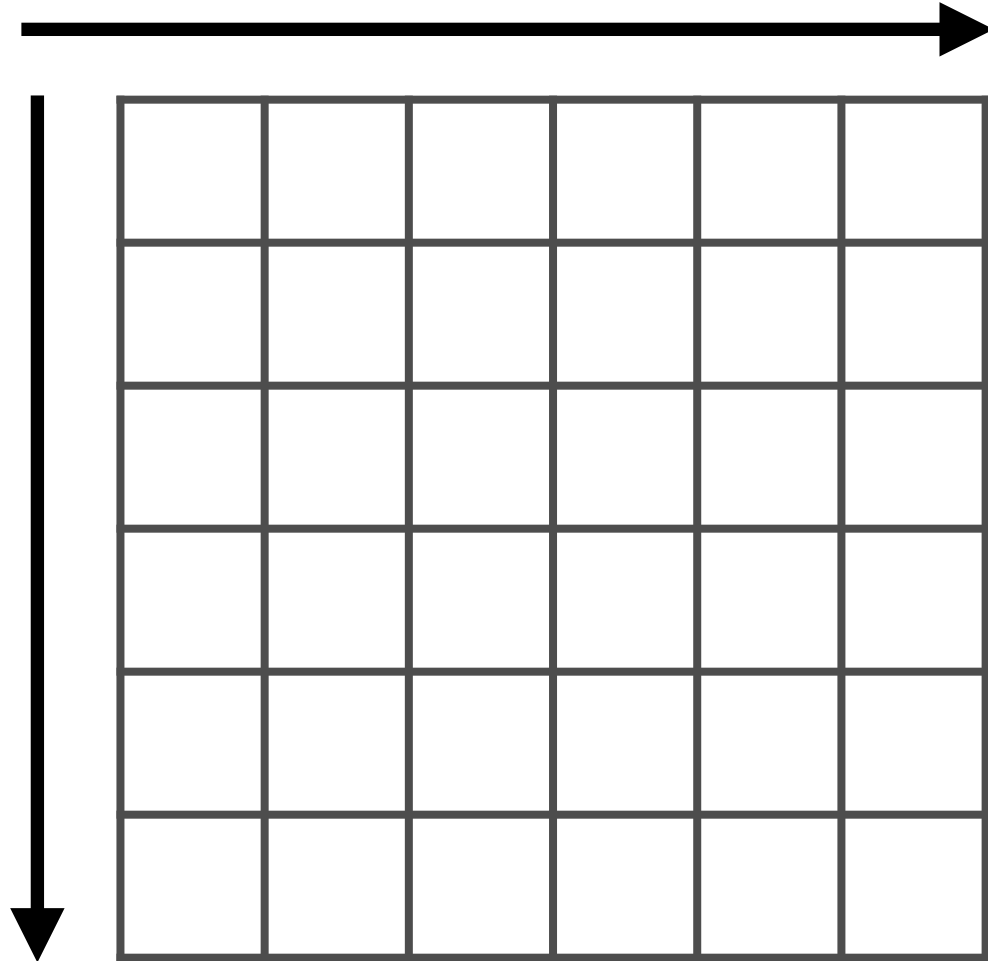**Each pixel represents only intensity information**

0.0 - 1.0
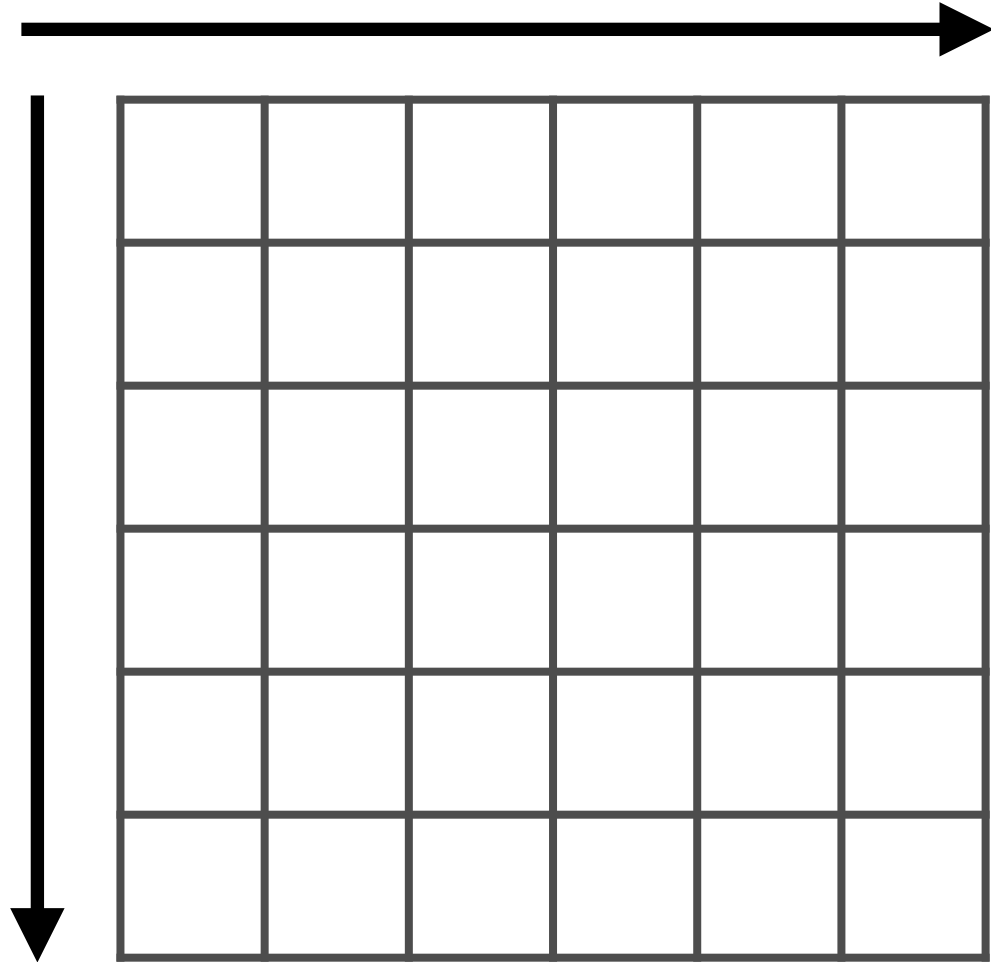
Grayscale Images

0.5

1 value to represent intensity, **1** channel

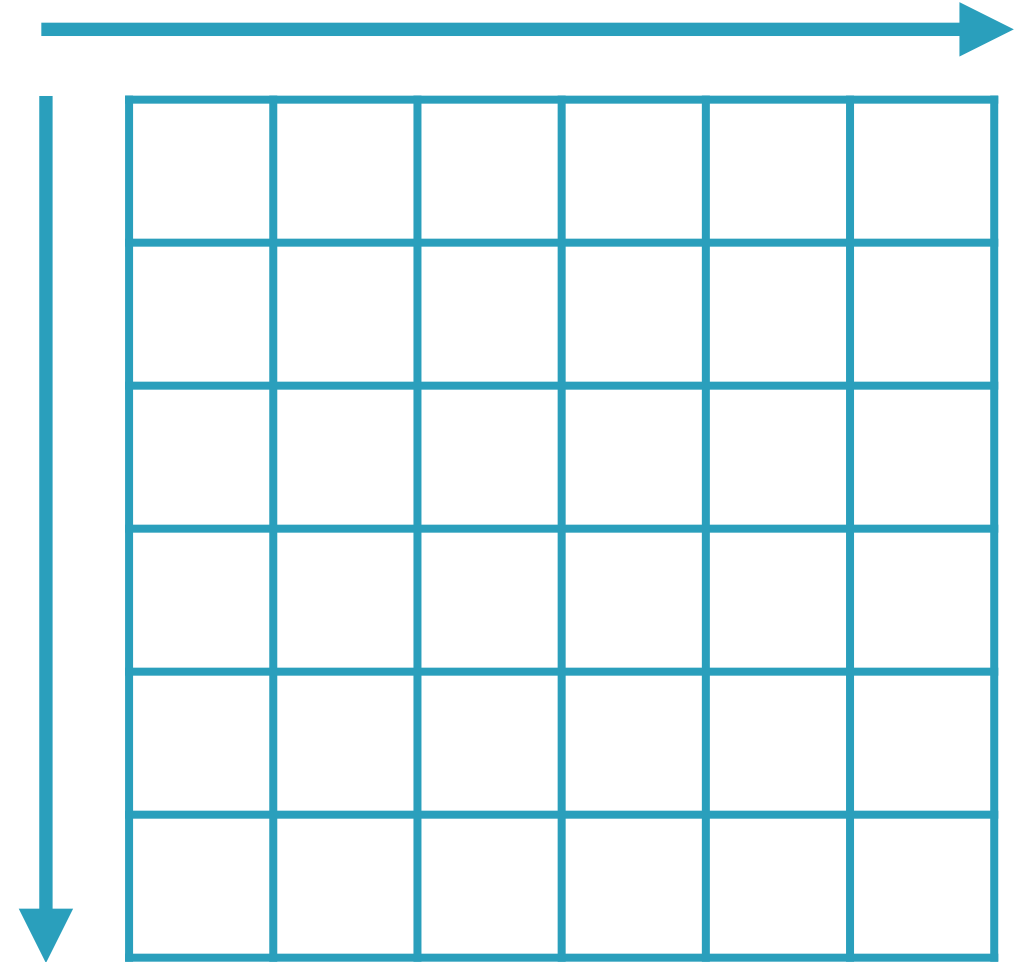# Images as Matrices

Images can be represented by a 3-D matrix

Images as Tensors

(6, 6, 1)   (6, 6, 3)

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

Performing classification on image data using neural networks in scikit-learn

Summary

Representing text as features

Term Frequency (TF) and Inverse Document Frequency (IDF)

Representing images as features

Single channel and multi-channel images

Text and image classification