Unsupervised Learning

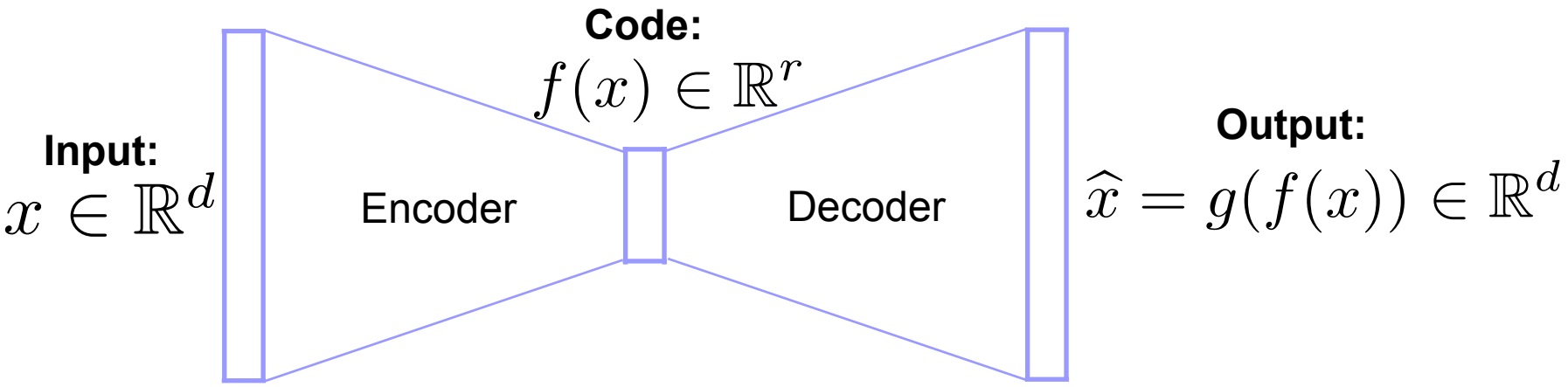
Machine Learning – CSE446

Kevin Jamieson
University of Washington

June 3, 2019

Autoencoders

Find a low dimensional representation for your data by predicting your data



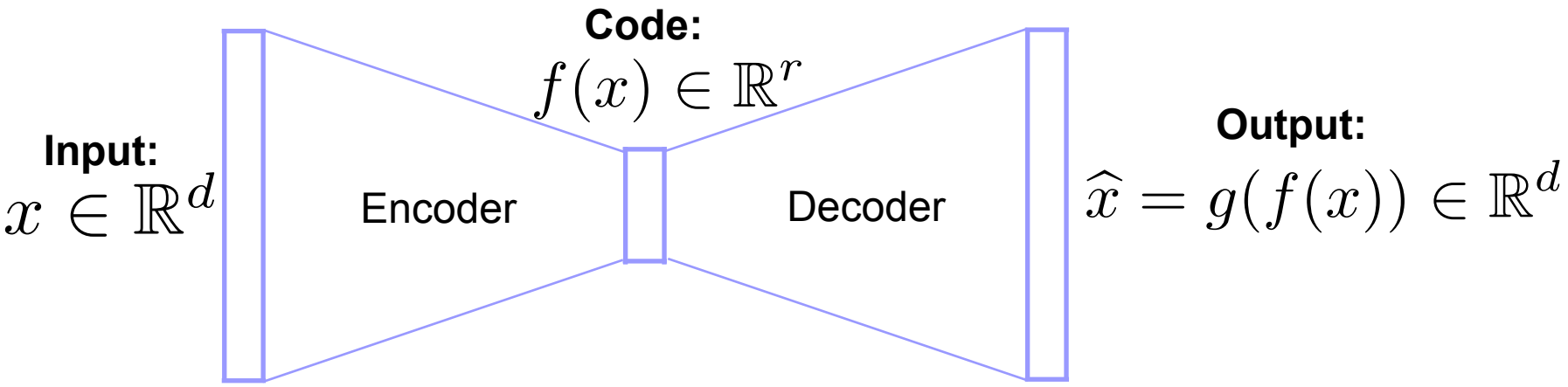
$$\underset{f,g}{\text{minimize}} \sum_{i=1}^{n} \|x_i - g(f(x_i))\|_2^2$$

PCA is a linear transformation

Three major applications
1. De-noising
2. Feature extraction
3. Manifold learning

Autoencoders

Find a low dimensional representation for your data by predicting your data

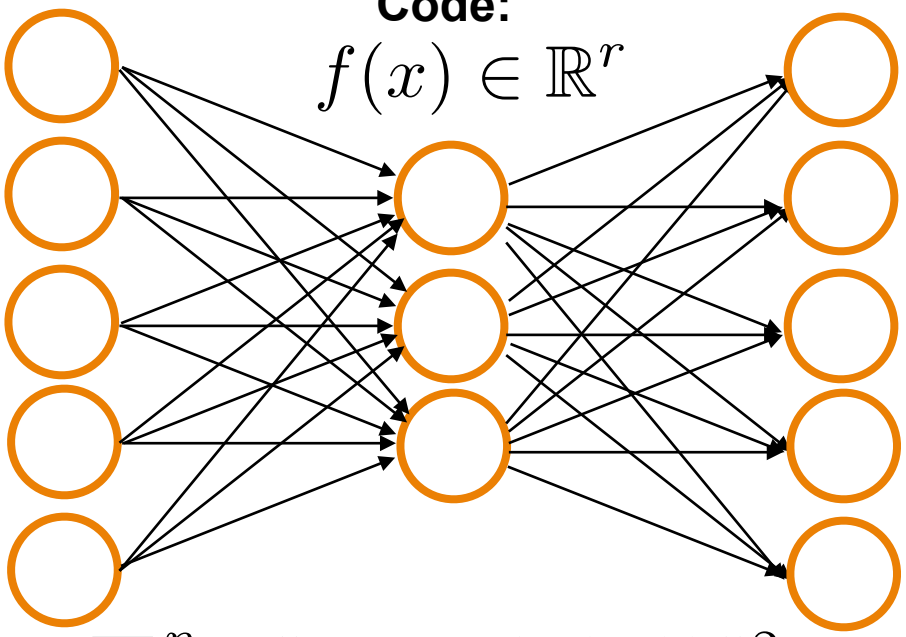


$$\underset{f,g}{\text{minimize}} \sum_{i=1}^{n} \|x_i - g(f(x_i))\|_2^2$$

Sparse Autoencoder

$$\underset{f,g}{\text{minimize}} \sum_{i=1}^{n} \|x_i - g(f(x_i))\|_2^2 + \lambda \|f(x_i)\|_1$$

Autoencoders

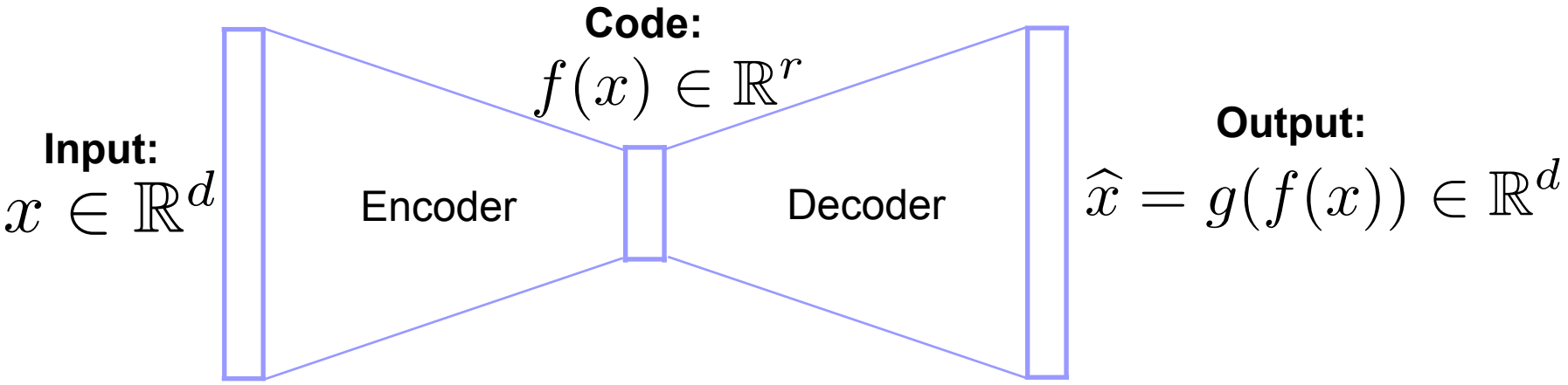


What if $f(X) = Ax$ and $g(y) = By$?

This will just pretty much be PCA…

Autoencoders

Find a low dimensional representation for your data by predicting your data



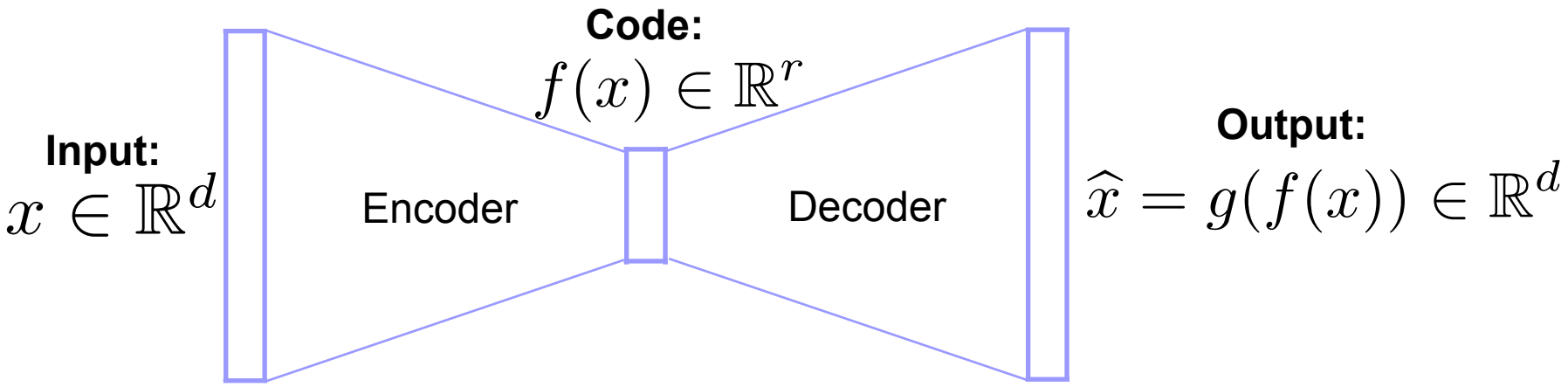
$$\underset{f,g}{\text{minimize}} \sum_{i=1}^{n} \|x_i - g(f(x_i))\|_2^2$$

Three major applications
1. De-noising
2. Feature extraction
3. Manifold learning

Related application: *Generating new samples (see autoencoder)*

GAN (Generative Adversarial Networks)

Find a low dimensional representation for your data by predicting your data



$$\underset{f,g}{\text{minimize}} \sum_{i=1}^{n} \|x_i - g(f(x_i))\|_2^2$$

Three major applications
1. De-noising
2. Feature extraction
3. Manifold learning

Related application: *Generating new samples* (see variational autoencoders (VAE) or generative adversarial networks (GAN)

Sequences and Recurrent Neural Networks

Machine Learning – CSE446

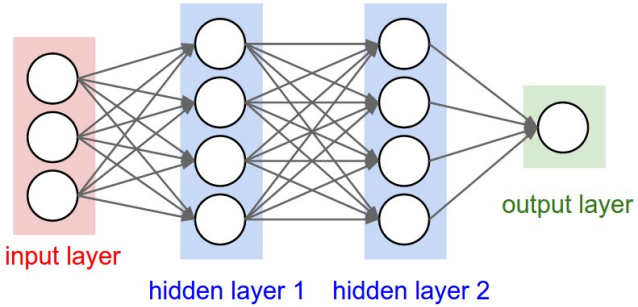
Kevin Jamieson
University of Washington

June 3, 2019

Variable length sequences

Images are usually standardized to be the same size (e.g., 256x256x3)

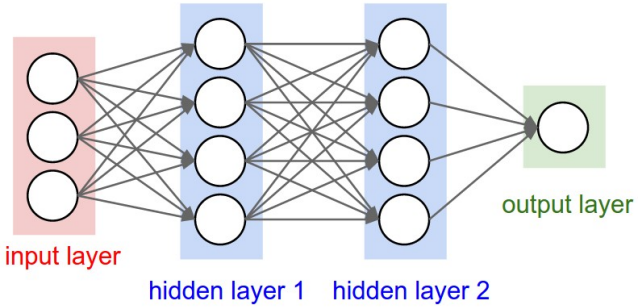
Neural Network



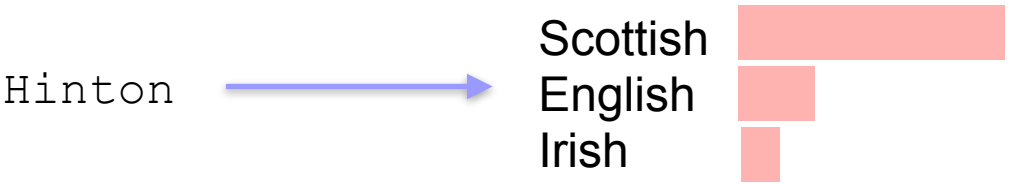
Variable length sequences

Images are usually standardized to be the same size (e.g., 256x256x3)

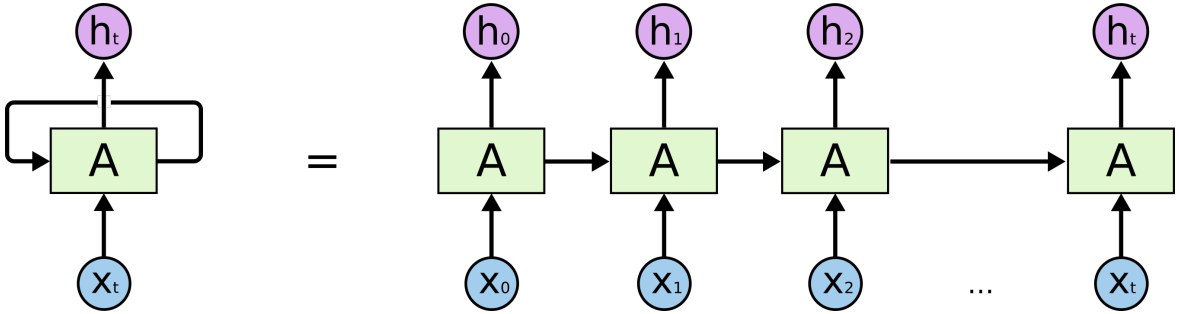
Neural Network



But what if we wanted to do classification on country-of-origin for names?

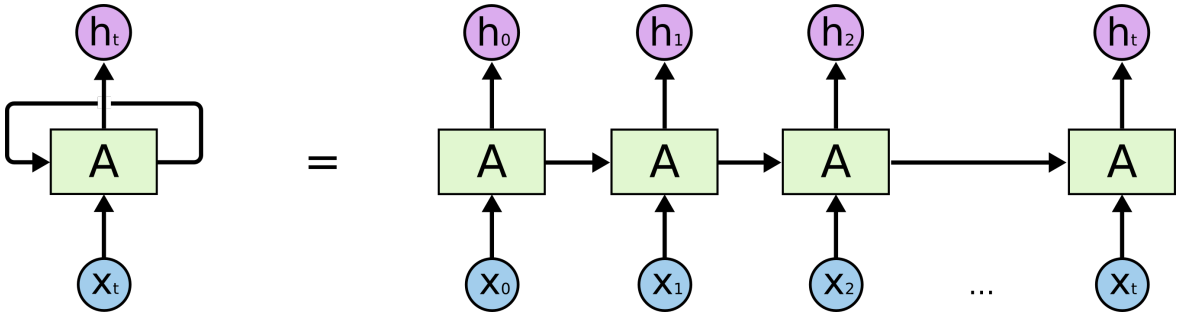


Recurrent Neural Network

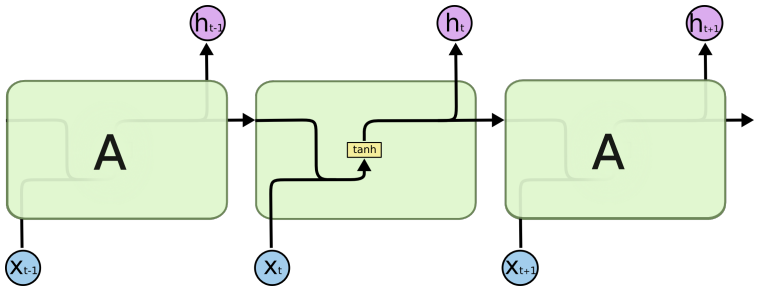


Variable length sequences

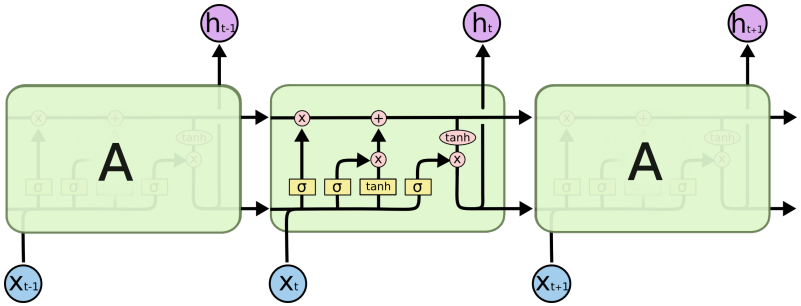
Recurrent Neural Network

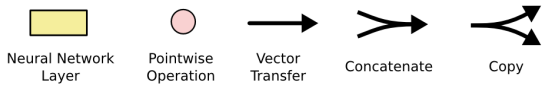


Standard RNN



Gated RNN
LSTM





Slide: http://colah.github.io/posts/2015-08-Understanding-LSTMs/