281 Live Session

Week 13 — 2023/4/12

Agenda

Non-Linear Classifiers

Basic CNN Architecture Types

Exercise – Face Classification Part 2

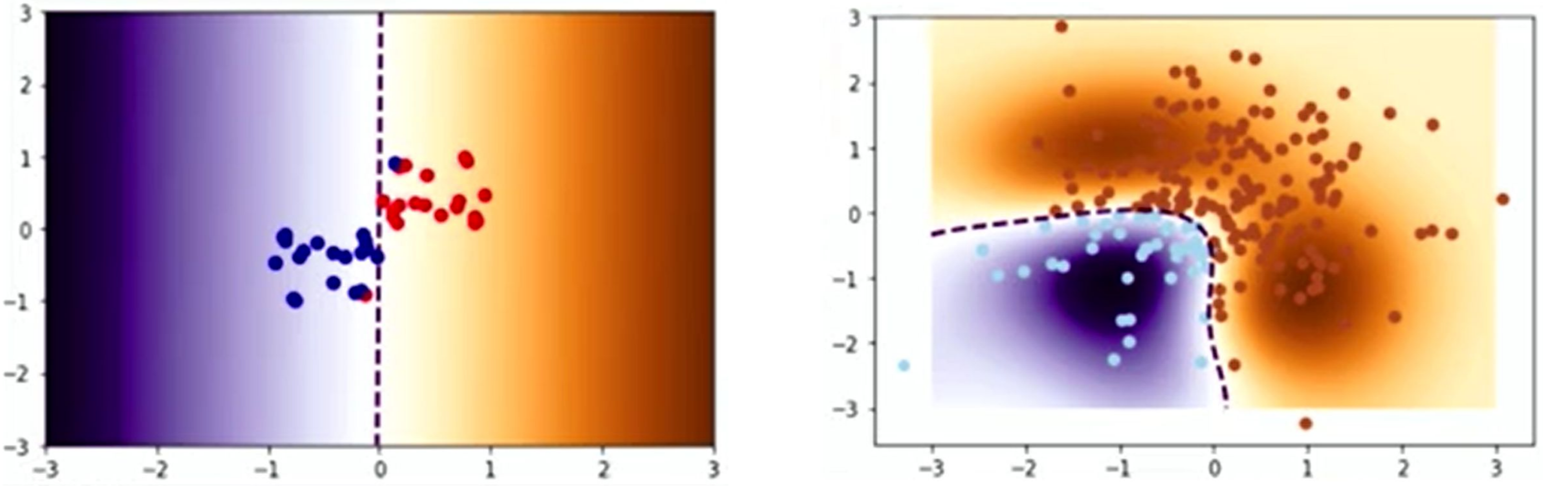
Modern Computer Vision Systems

Non-Linear Classifiers

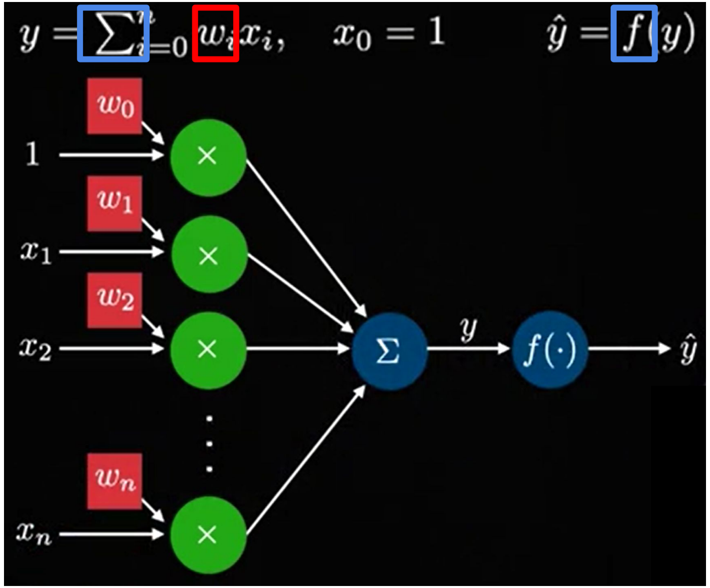
Intuition goals:

- Differences between linear and nonlinear SVMs
- What is a neuron? What functions can it perform?
- What is changing in a network on each iteration?
- What is an activation function? How do they work?

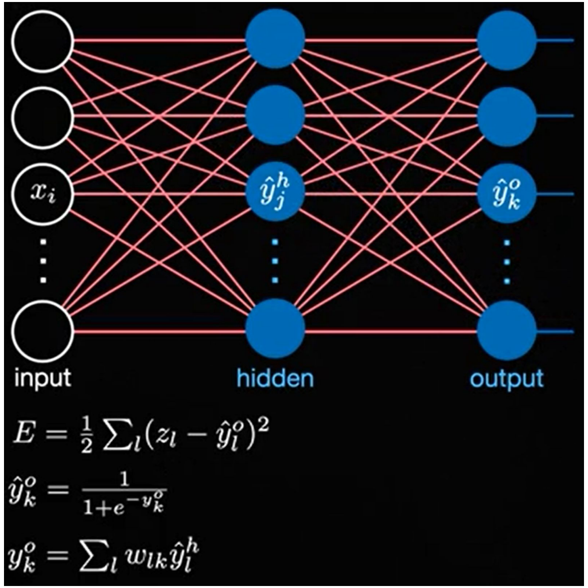
Linear vs Nonlinear SVMs



Neural Networks



x → data
z → labels
w → weights (learned)
f → activation function
α → step size
Δ → direction of gradient



$$E(\vec{w}) = \sum_{k=1}^{m}(z_k - f(\vec{x}_k^T \cdot \vec{w}))^2$$

$$\Delta w_i = \sum_{k=1}^{m} \alpha(z_k - f(\vec{x}_k^T \cdot \vec{w}))(x_{k,i})(f'(\vec{x}_k^T \cdot \vec{w}))$$

Non-Linear Classifiers

Intuition goals:

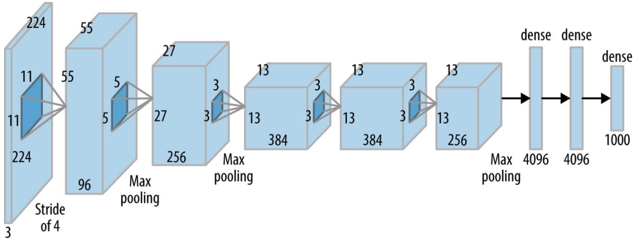
- Differences between linear and nonlinear SVMs
- What is a neuron? What functions can it perform?
- What is changing in a network on each iteration?
- What is an activation function? How do they work?

Model Building Steps

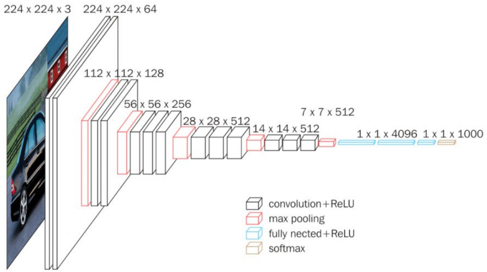
1. Inspect your data
   - Can it be standardized? Cleaned up? Augmented?
2. Split data into training, test, and validation set
3. Choose features to extract
   - What information do you need for training?
   - Are your features robust across all images?
4. Decide on a model
   - Linear? Nonlinear? Neural?
5. Define the error function
   - Usually quadratic
6. Define the gradient function
   - Usually the derivative of the error
7. Test a range of hyperparameters
8. Iterate!

AlexNet, VGGNet, & ResNet

AlexNet (original deep CNN)



VGGNet (fewer parameters)



ResNet (skip connections)

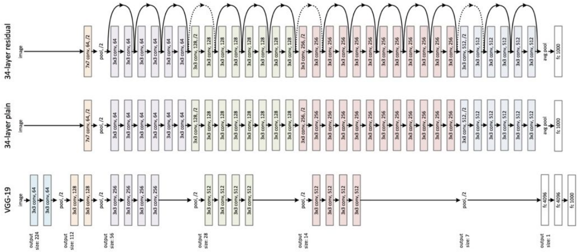


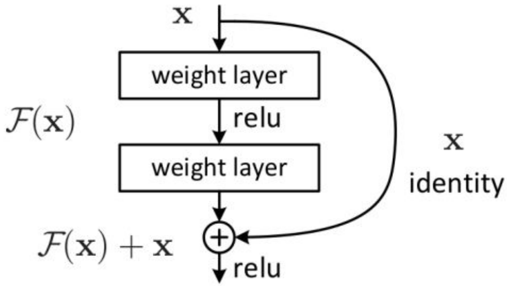
Figure 3. Example network architectures for ImageNet. **Left**: the VGG-19 model [41] (19.6 billion FLOPs) as a reference. **Middle**: a plain network with 34 parameter layers (3.6 billion FLOPs). **Right**: a residual network with 34 parameter layers (3.6 billion FLOPs). The dotted shortcuts increase dimensions. **Table 1** shows more details and other variants.

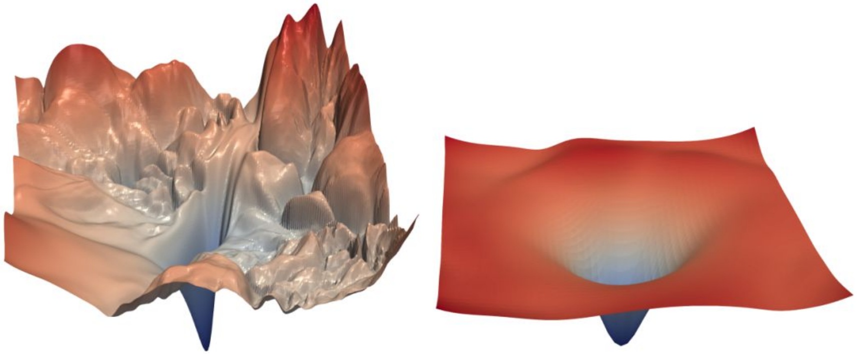
source: towardsdatascience.com, analyticsvidhya.com

Benefits of Skip Connections

Addresses vanishing gradient problem

Can be combined using addition or concatenation





(a) without skip connections

(b) with skip connections

source: towardsdatascience.com, analyticsvidhya.com

Exercise — Faces Part 2 (ResNet Features)



Modern Computer Vision Systems

Generative Adversarial Networks (GANs)

- Unsupervised learning (look ma, no labels!)
- Simultaneously learn the *generator* (makes images) and the *discriminator* (decides if images are fake or real)

Unsolved problems for GANs

- Unstable / hard to train reliably
- Not usually good at long-range correlations (eg earrings)

Source: thispersondoesnotexist.com, github.com/NVlabs/stylegan2

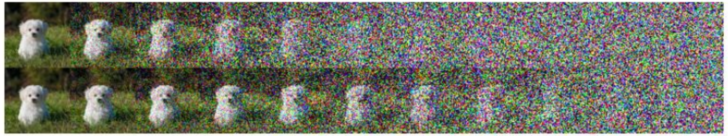
Modern Computer Vision Systems

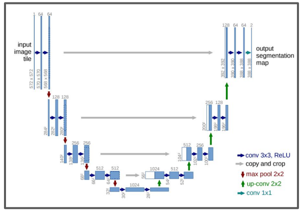
Diffusion Models

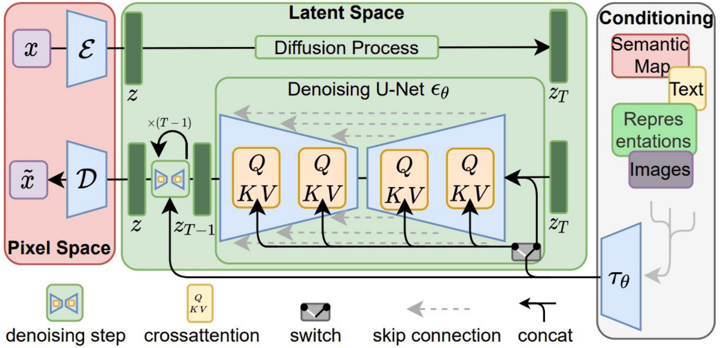
- Similar to an auto-encoder, except using noise instead of down-sampling
- Can be used directly for denoising/superresolution
- Or, can be guided for image generation

Unsolved problems for Diffusion Models

- Structure vs texture limitations
- Difficult to control







Source: theaisummer.com/diffusion-models, en.wikipedia.org/wiki/DALL-E

Modern Computer Vision Systems

There is still significant work needed to improve the usefulness of these systems

- Robustness
- Data privacy
- Bias / ethics concerns
- Explainability

Upcoming ToDo's

Watch async lectures for unit 13

Course evaluations next week

Final project presentations next week (10 minutes!)

Written report due April 23rd