Generative models

Neural Networks Design And Application

Machine learning paradigm

Training data

ML model for
panda recognition

Testing data

Machine learning paradigm

x

Training data

ML model for
panda recognition

Testing data

Machine learning paradigm

ML model for
panda recognition

Testing data

x → $P(y|x)$

Training data

Machine learning paradigm

ML model for panda recognition

Testing data

x

$P(y|x)$

$P(+1|x)$

Panda

$P(-1|x)$

Not panda

Training data

Machine learning paradigm

ML model for panda recognition

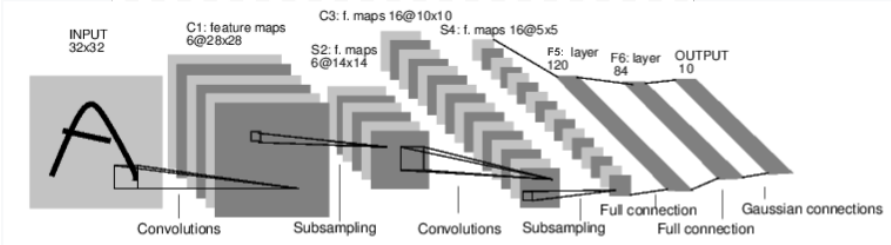
Testing data

x → mapping → $P(y|x)$

$P(+1|x)$ Panda

$P(-1|x)$ Not panda

Training data



Machine learning paradigm

ML model for panda recognition

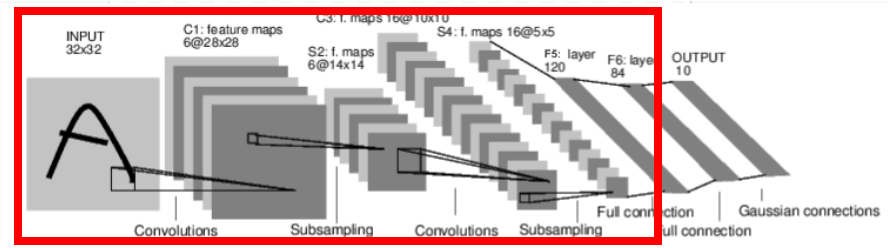
Testing data

x → mapping → $P(y|x)$

$P(+1|x)$ Panda

$P(-1|x)$ Not panda

Training data



Machine learning paradigm

ML model for panda recognition

Testing data

x

mapping

$P(y|x)$

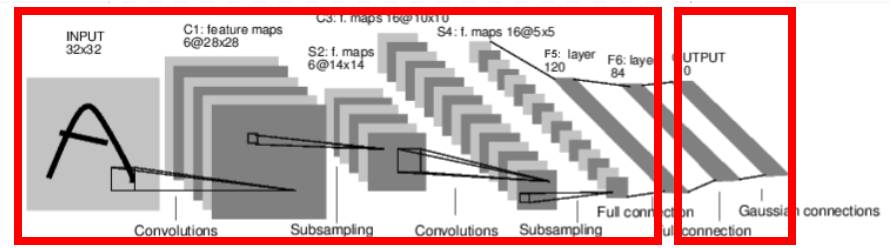
Panda

$P(+1|x)$

$P(-1|x)$

Not panda

Training data



Machine learning paradigm

ML model for panda recognition

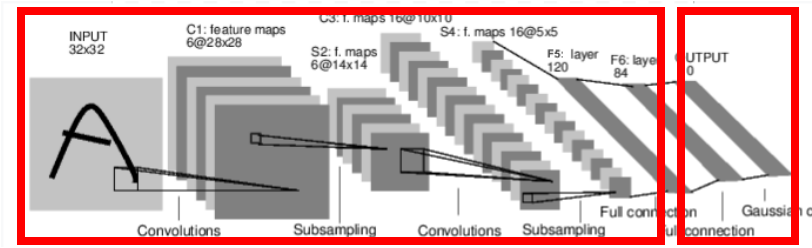
Testing data

x → mapping → $P(y|x)$

$P(+1|x)$ Panda

$P(-1|x)$ Not panda

Training data



$P(+1|x) = 0.97$

$P(-1|x) = 0.03$

Machine learning paradigm

Discriminative model

ML model for panda recognition

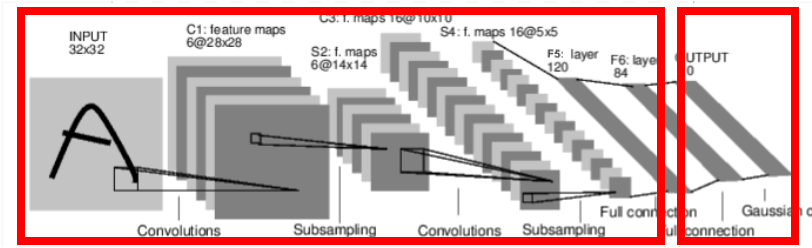
Testing data

x → mapping → $P(y|x)$

$P(+1|x)$ Panda

$P(-1|x)$ Not panda

Training data



$P(+1|x) = 0.97$

$P(-1|x) = 0.03$

Not panda

Machine learning paradigm

Q: other way for classification?

Discriminative model

ML model for panda recognition

Testing data

x → mapping → $P(y|x)$

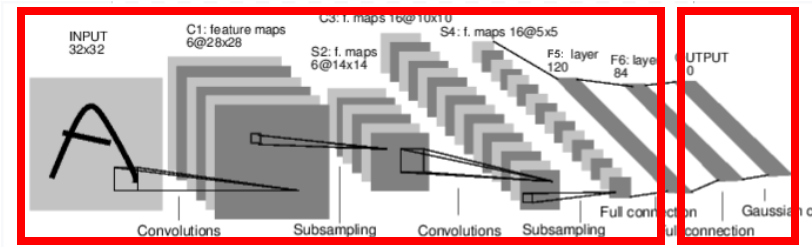
Panda

$P(+1|x)$

$P(-1|x)$

Not panda

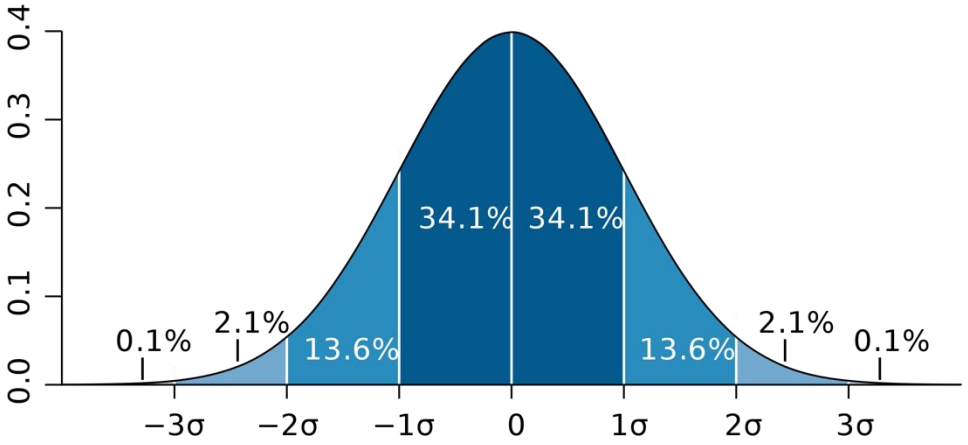
Training data



$P(+1|x) = 0.97$

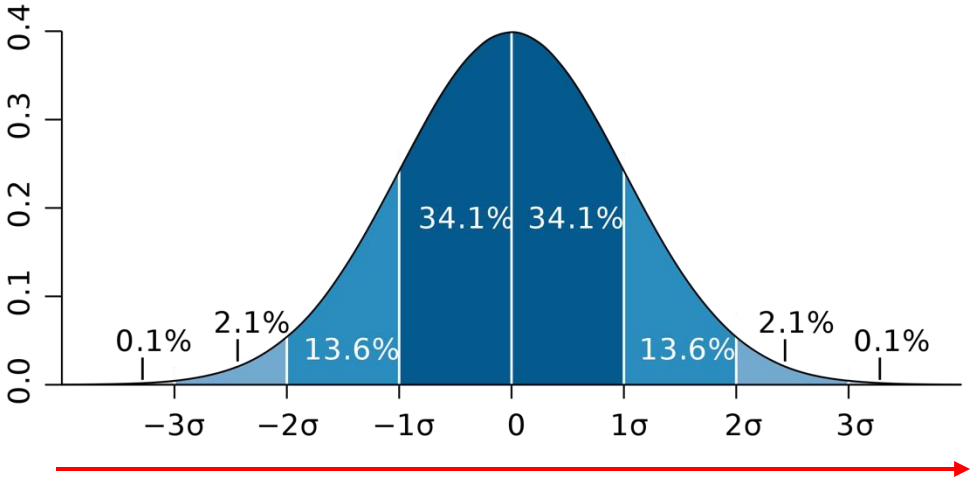
$P(-1|x) = 0.03$

What is a probability distribution?



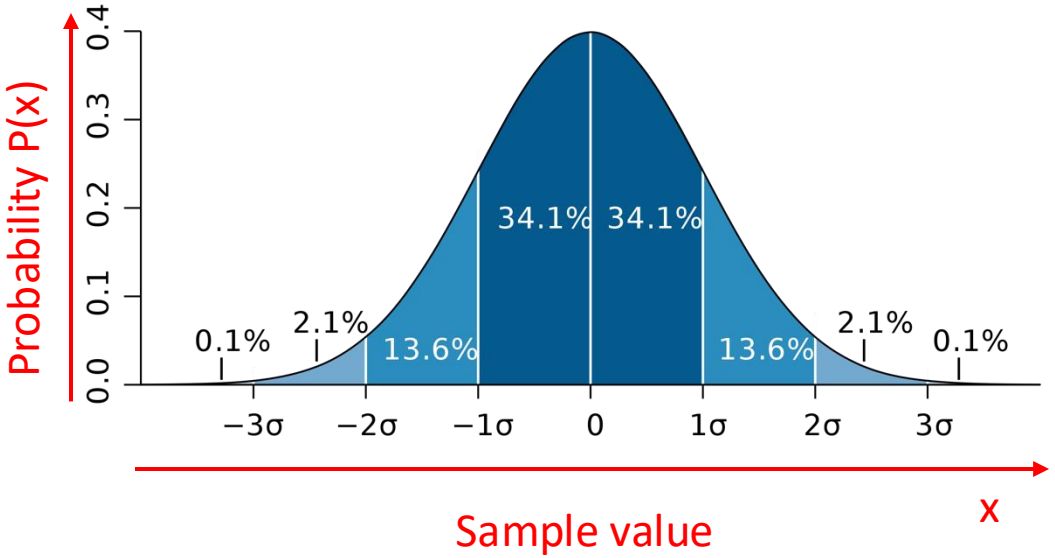
Images from https://en.wikipedia.org/wiki/Probability_distribution

What is a probability distribution?



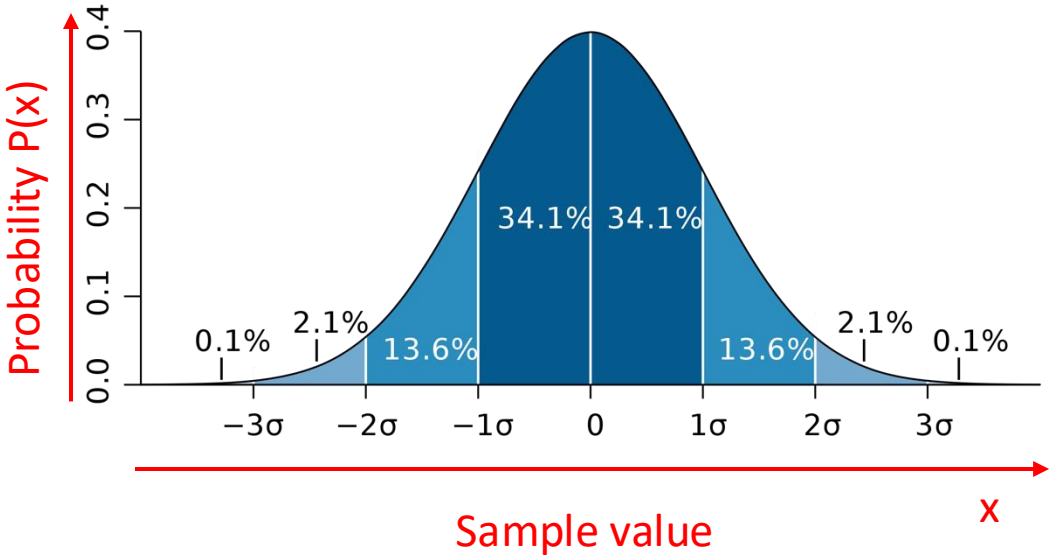
Images from https://en.wikipedia.org/wiki/Probability_distribution

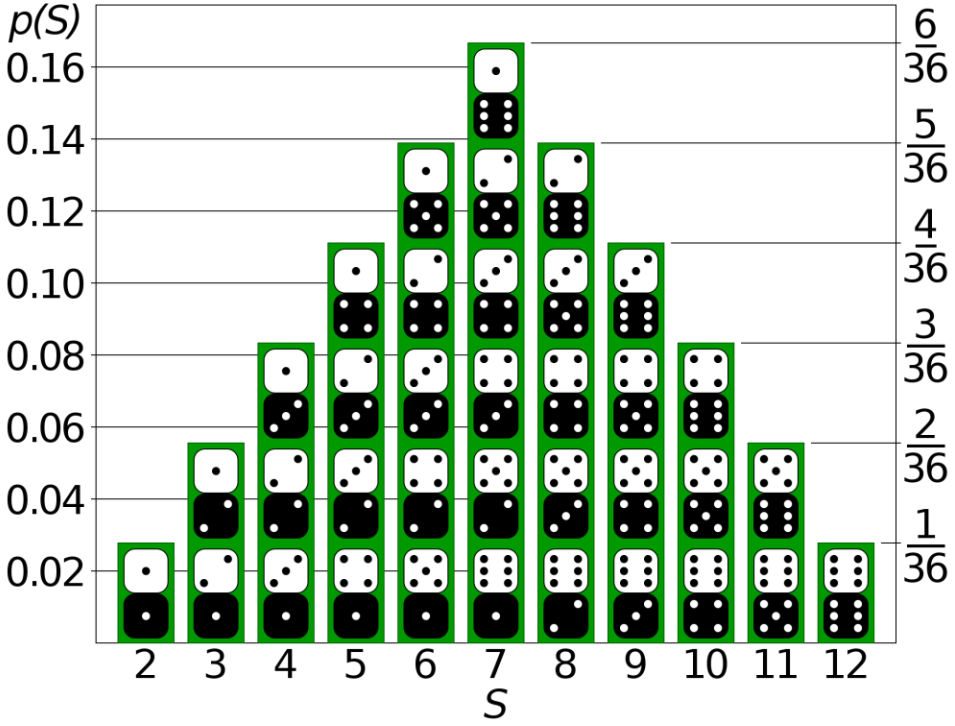
What is a probability distribution?



Images from https://en.wikipedia.org/wiki/Probability_distribution

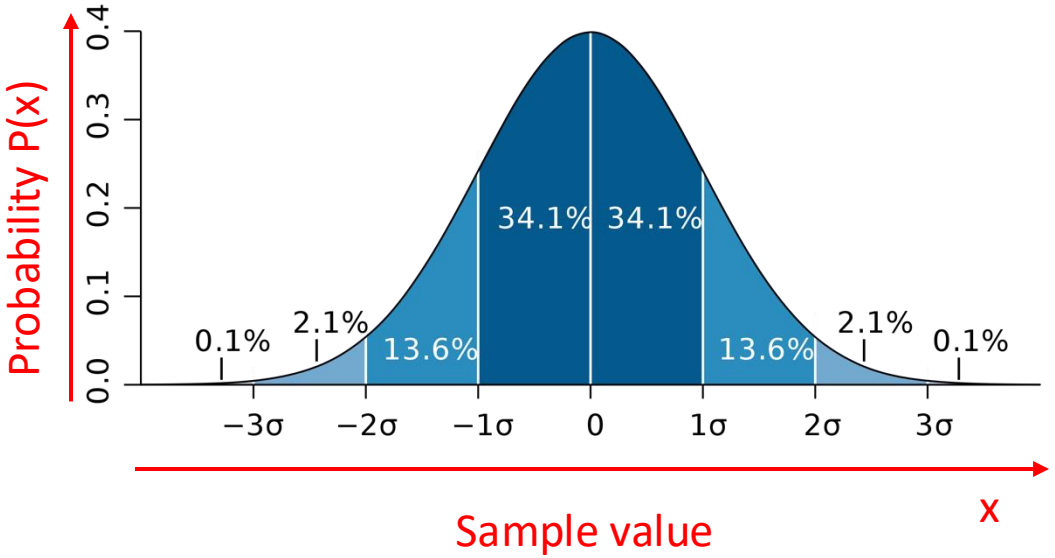
What is a probability distribution?

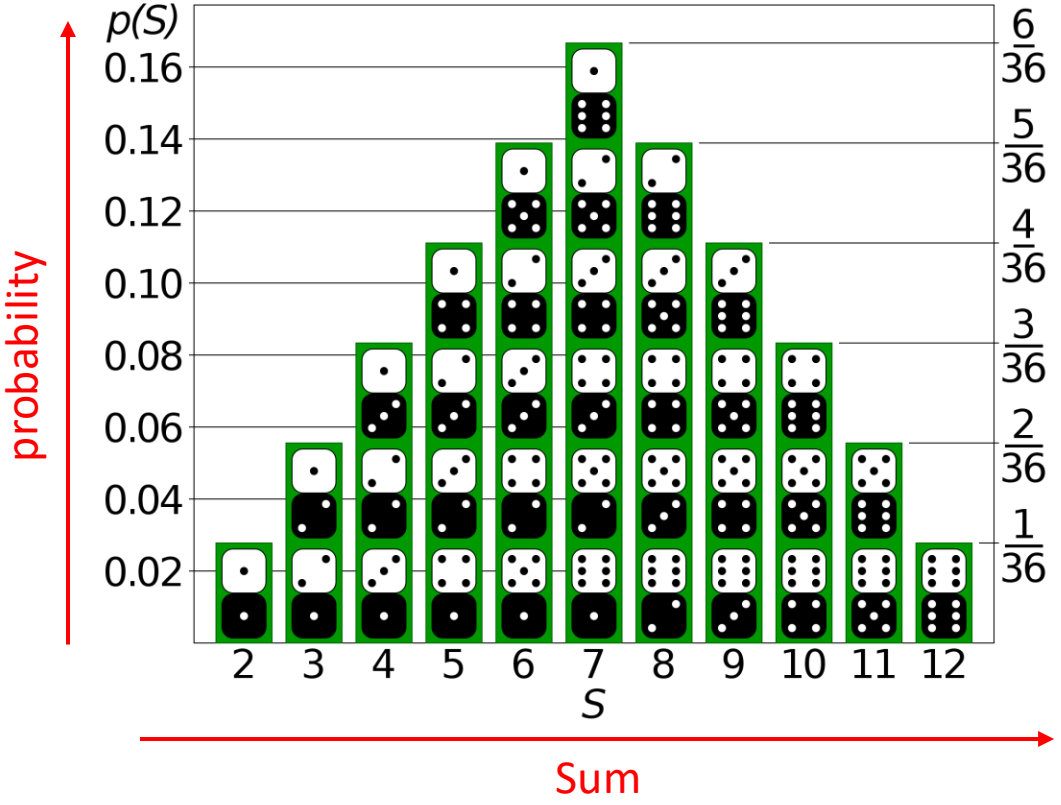




Images from https://en.wikipedia.org/wiki/Probability_distribution

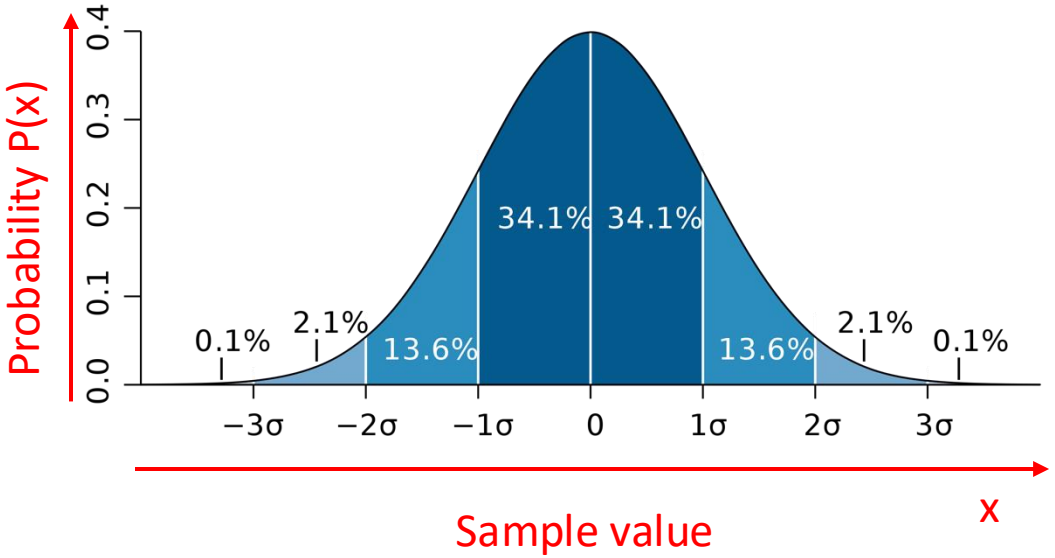
What is a probability distribution?



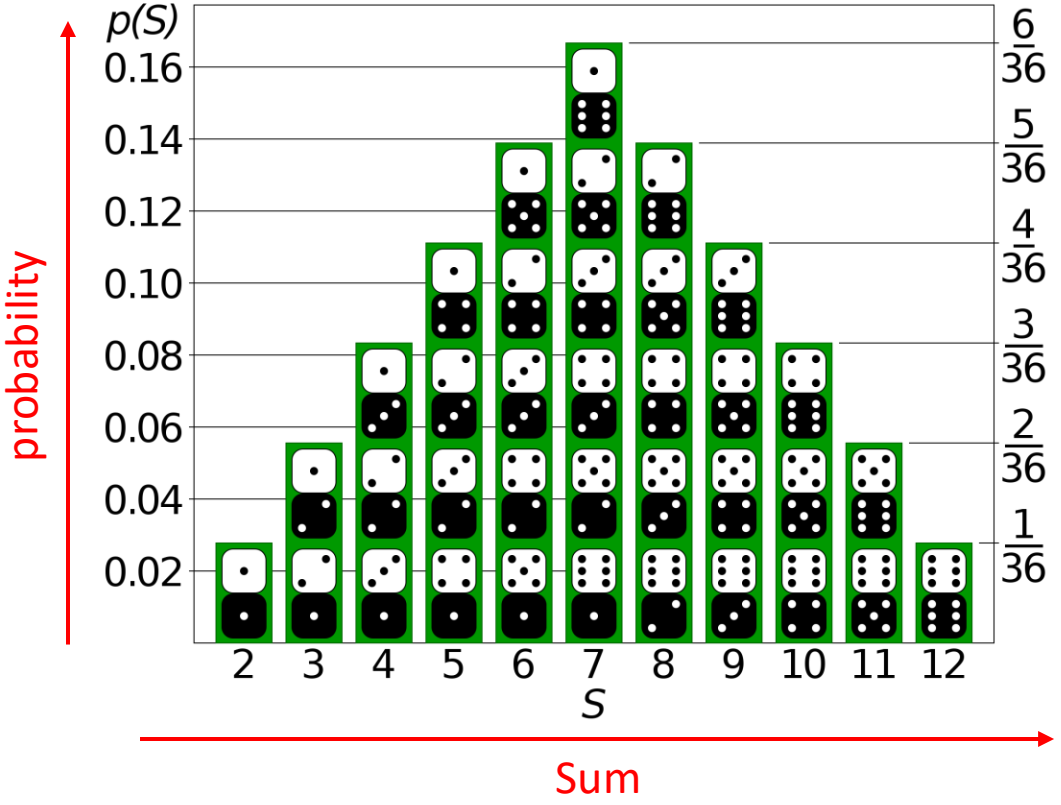


Images from https://en.wikipedia.org/wiki/Probability_distribution

What is a probability distribution?

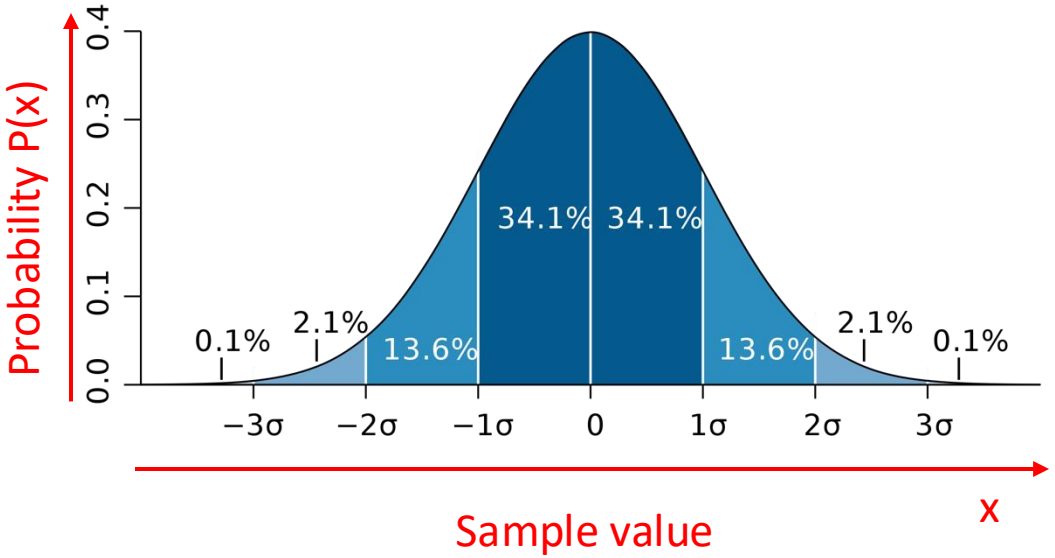


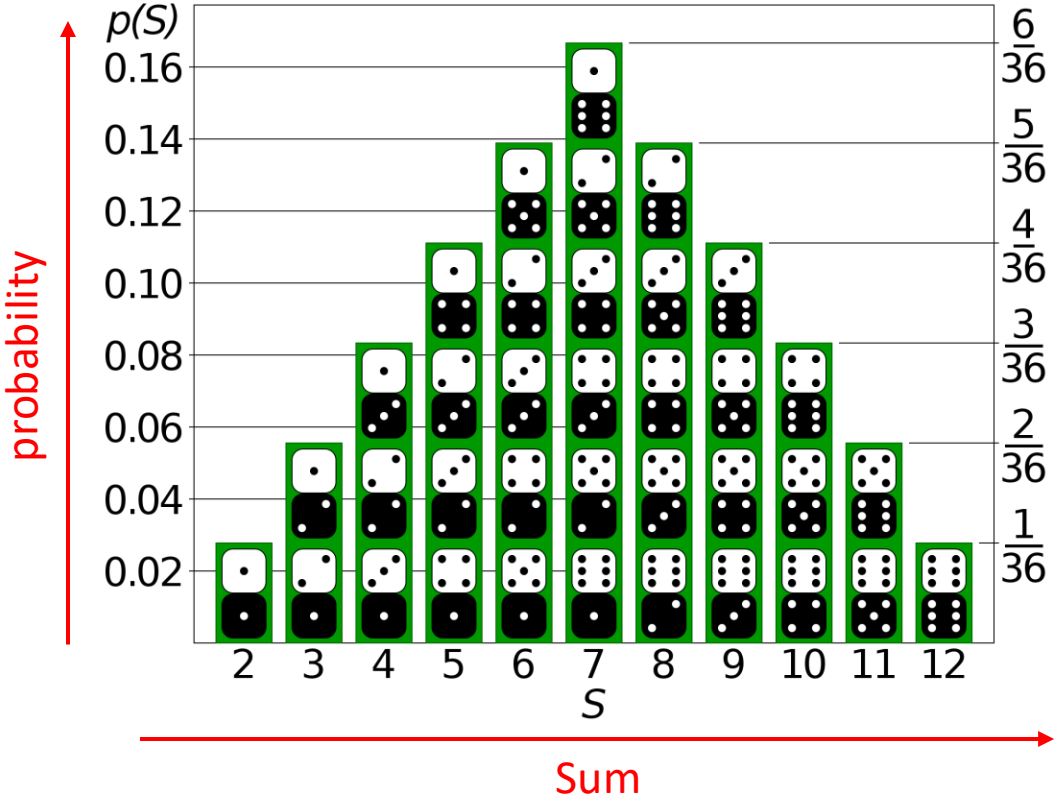
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



Images from https://en.wikipedia.org/wiki/Probability_distribution

What is a probability distribution?



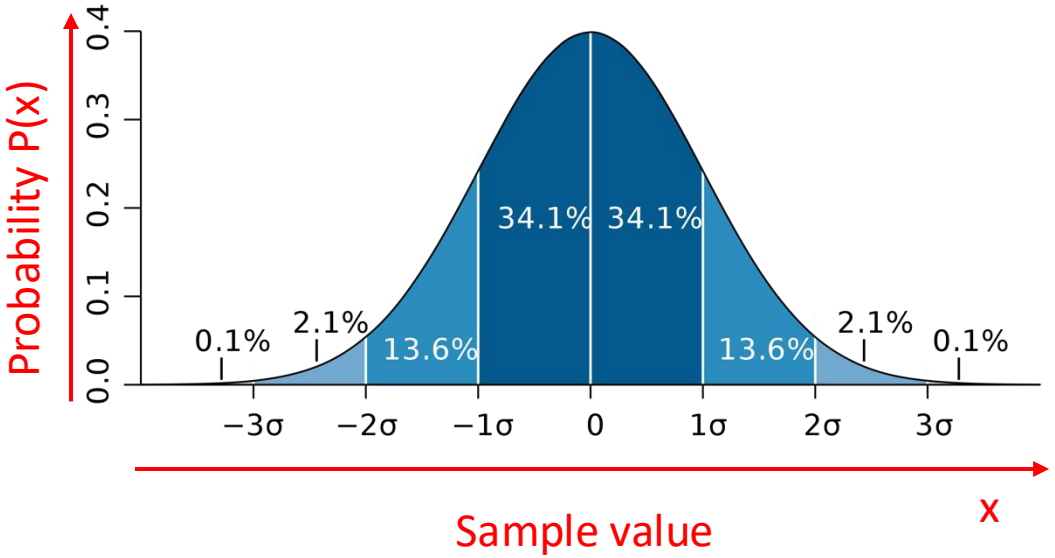


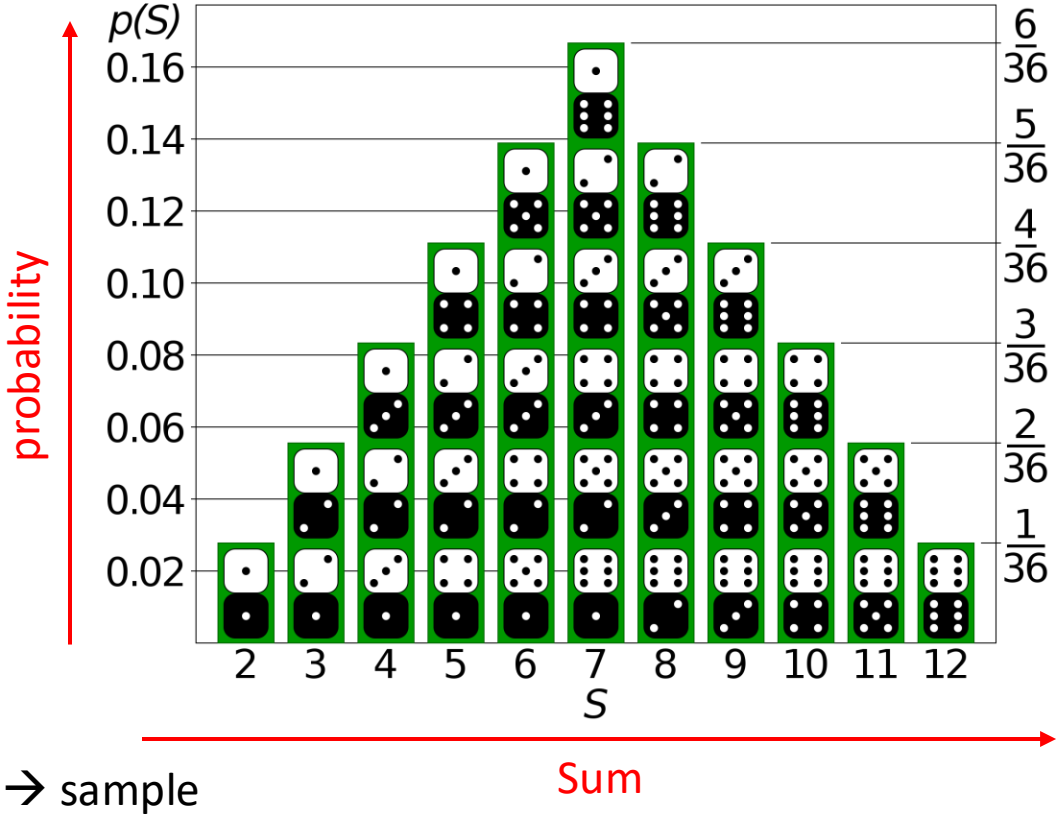
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \longrightarrow f^{-1}$$

Inverse distribution function

Images from https://en.wikipedia.org/wiki/Probability_distribution

What is a probability distribution?



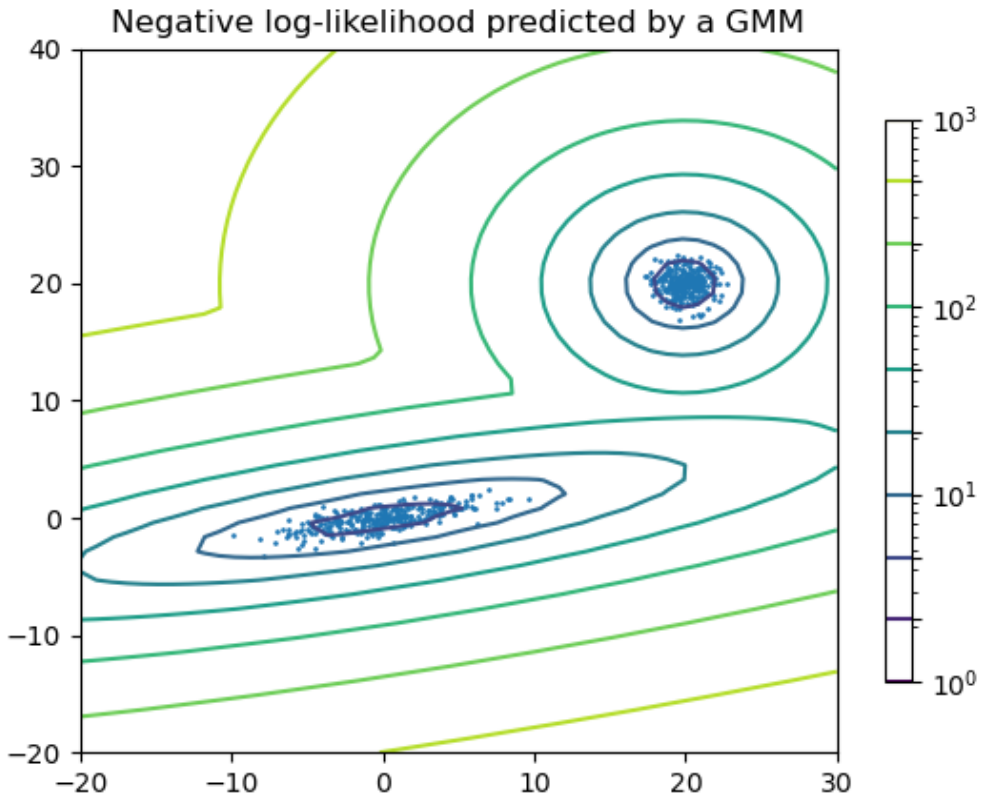


$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \longrightarrow f^{-1} \longrightarrow$$ Probability → sample

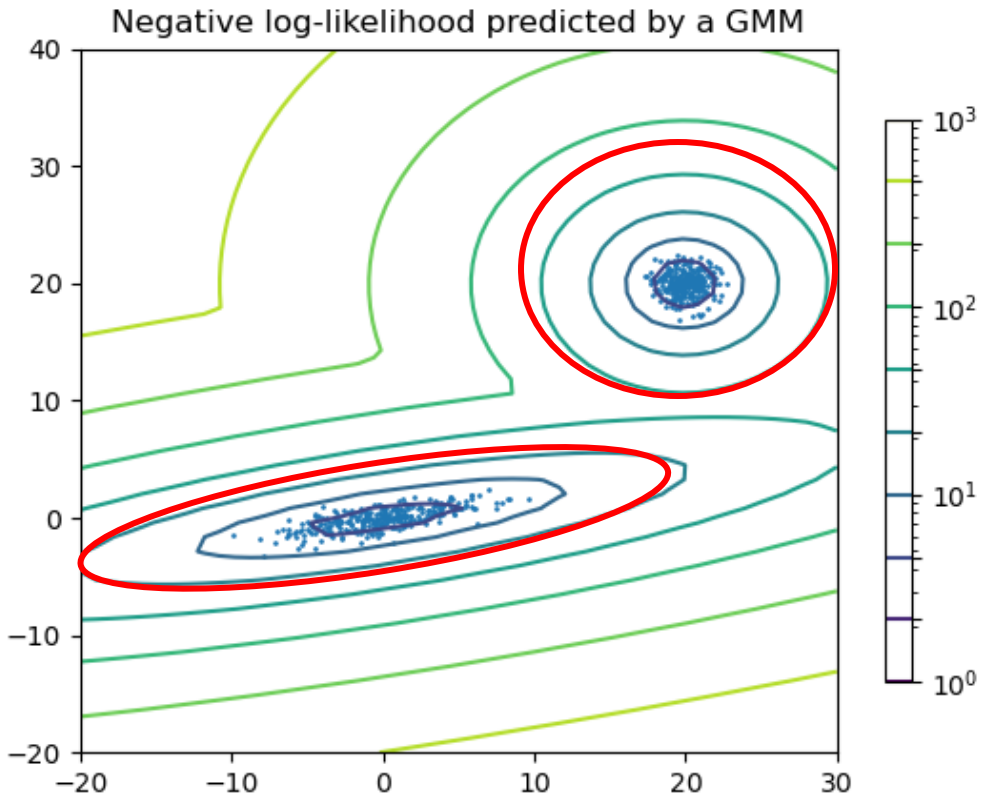
Inverse distribution function

Images from https://en.wikipedia.org/wiki/Probability_distribution

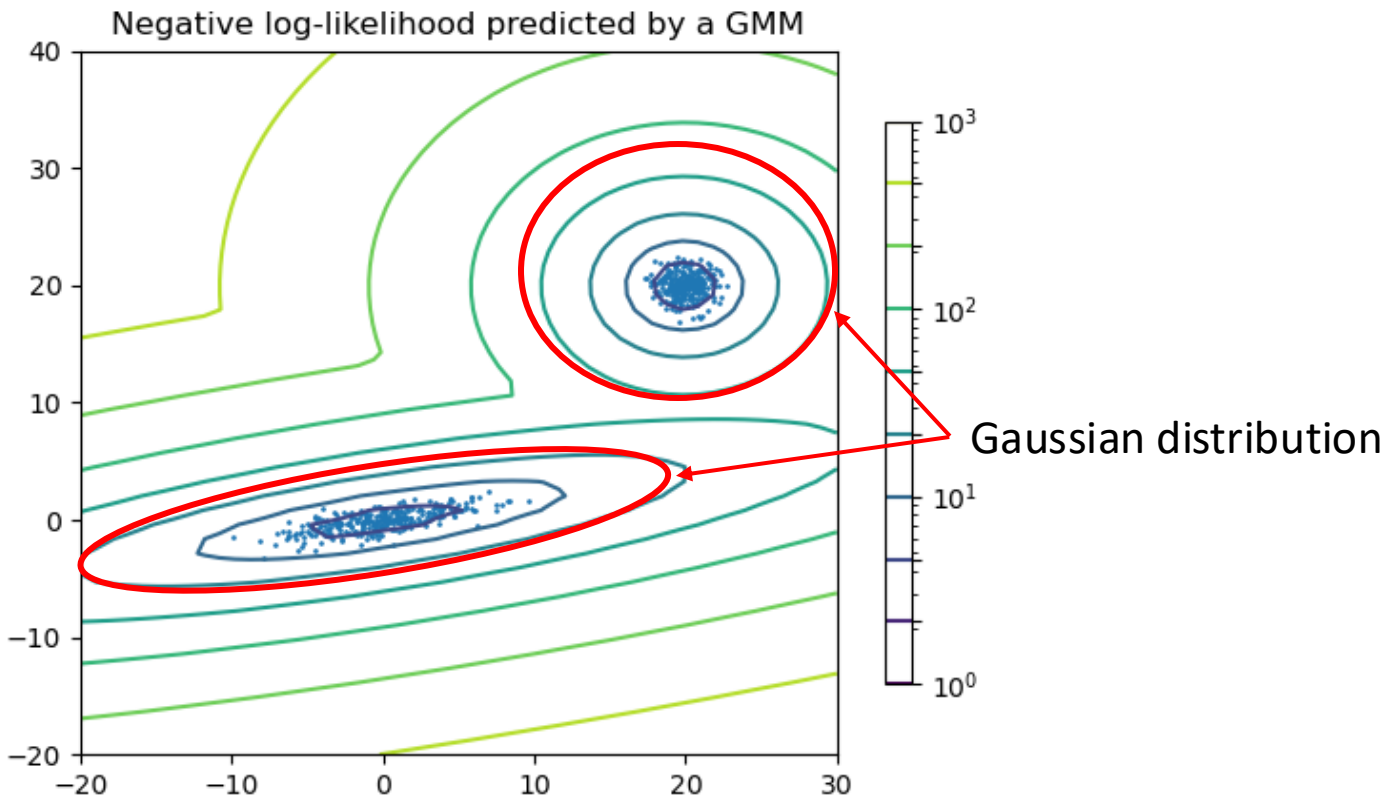
Gaussian mixture model



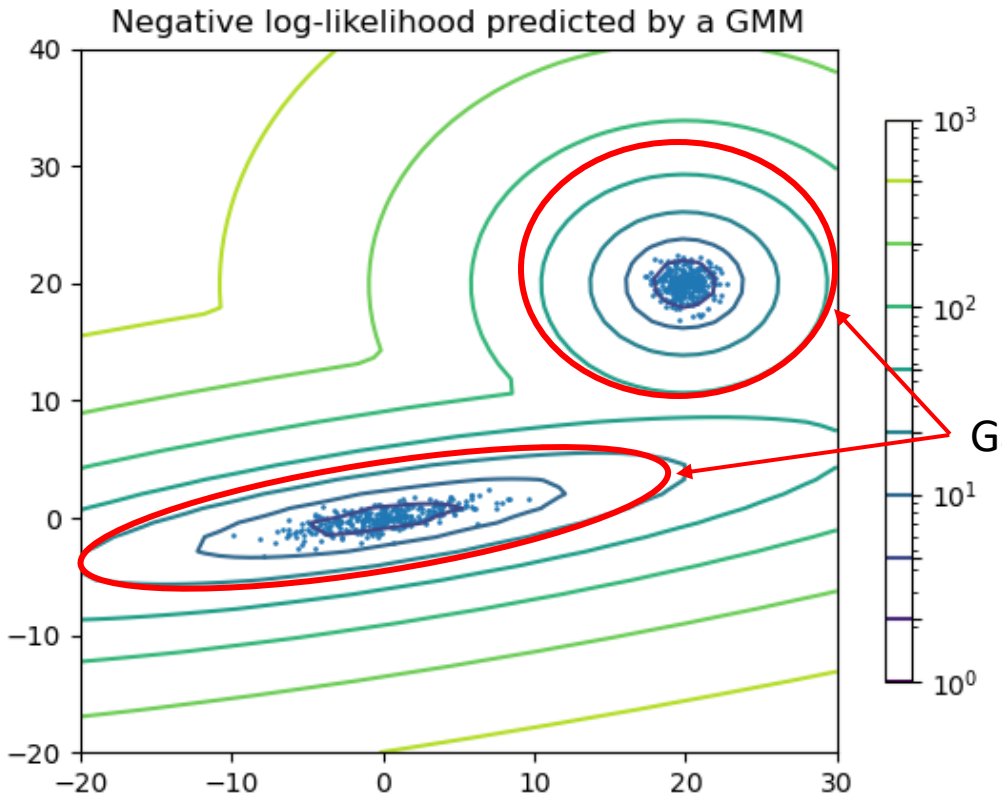
Gaussian mixture model



Gaussian mixture model



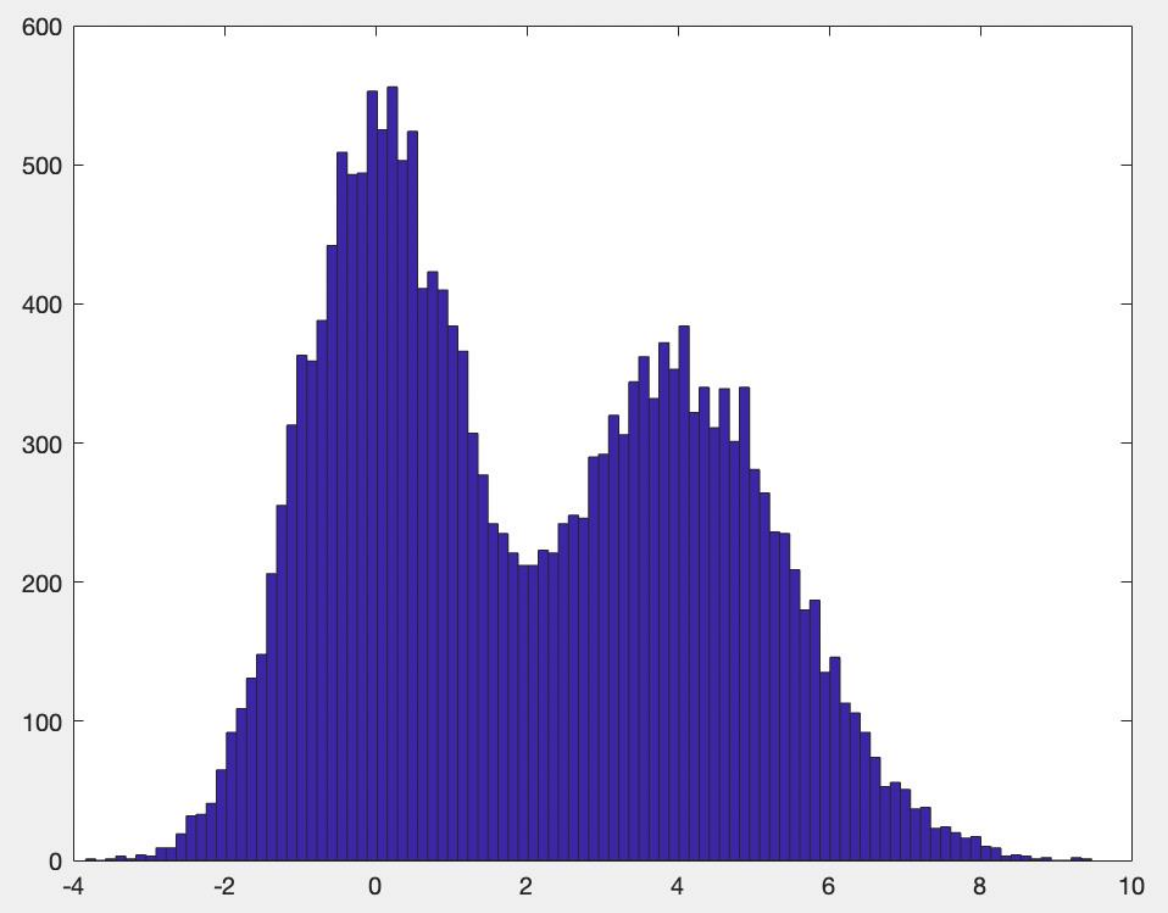
Gaussian mixture model



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

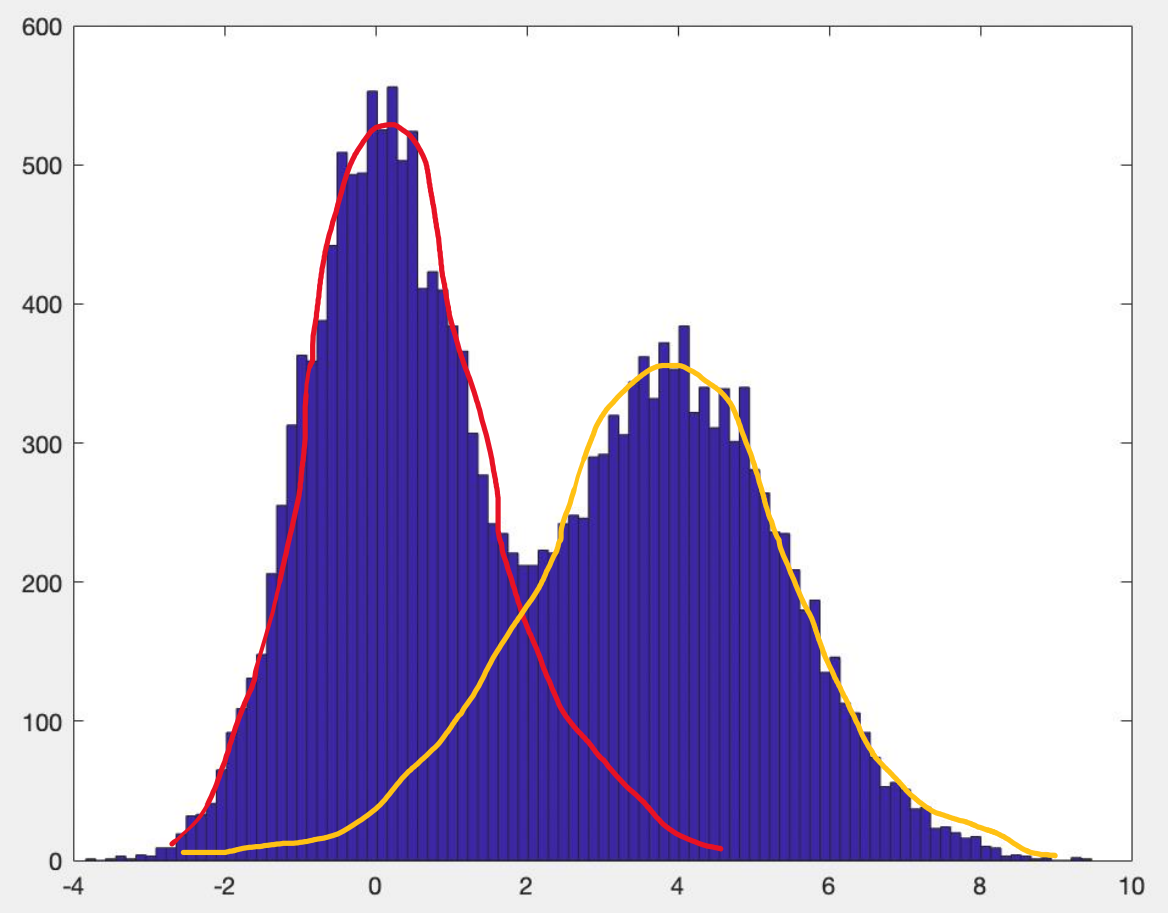
Gaussian distribution

Gaussian mixture model



Q: how many Gaussian distributions?

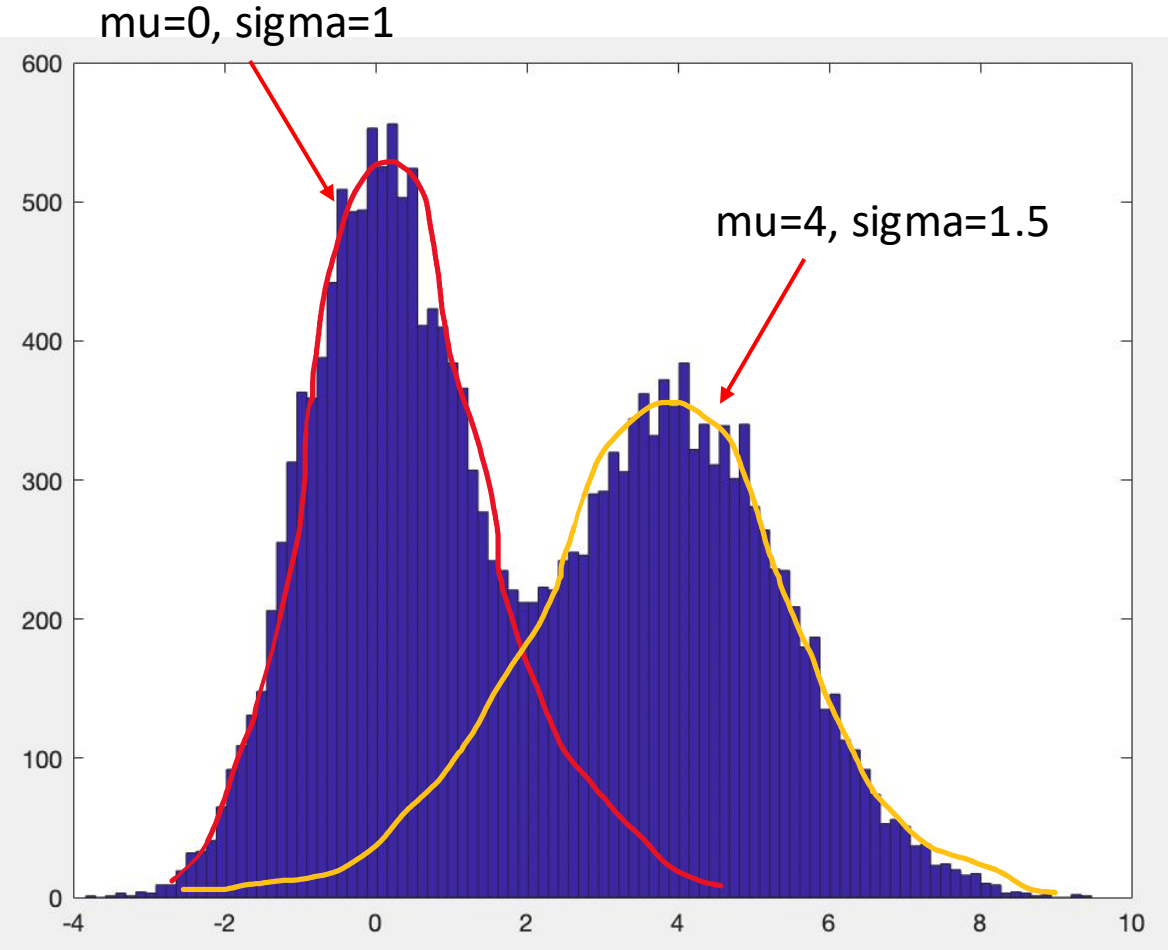
Gaussian mixture model



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Q: how many Gaussian distributions?

Gaussian mixture model



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Q: how many Gaussian distributions?

Machine learning paradigm

ML model for panda recognition

Testing data

x

$P(y|x)$

$P(+1|x)$

Panda

$P(-1|x)$

Not panda

Training data

Machine learning paradigm

ML model for panda recognition

Testing data

x

$P(y|x)$

$P(+1|x)$

Panda

$P(-1|x)$

Not panda

Training data

Machine learning paradigm

ML model for panda recognition

Testing data

x

Learning f (f^{-1})

$P(y|x)$

$P(x|y)$

Training data

Machine learning paradigm

ML model for panda recognition

Testing data

x → Learning f (f^{-1}) → $P(y|x)$

Training data

Generative models ← $P(x|y)$

Machine learning paradigm

ML model for panda recognition

Testing data

x

$P(x|x)$

Learning f (f^{-1})

$P(x|y)$

Training data

Generative models

Q: how to classify a new data?

Machine learning paradigm

$$P(x', y) = P(x'|y)P(y)$$

ML model for panda recognition

Testing data

x → Learning f (f^{-1}) → ~~$P(x|x)$~~ $P(x|y)$

Training data

Generative models

Q: how to classify a new data?

Machine learning paradigm

$$P(x', y) = P(x'|y)P(y)$$

Pick the label y with largest $P(x', y)$

ML model for panda recognition

Testing data

x

Learning f (f^{-1})

~~$P(x|x)$~~

$P(x|y)$

Training data

Generative models

Q: how to classify a new data?

Machine learning paradigm

$$P(x', +1) = P(x'|+1)P(+1)$$
$$P(x', -1) = P(x'|-1)P(-1)$$

$$P(x', y) = P(x'|y)P(y)$$

Pick the label y with largest $P(x', y)$

ML model for panda recognition

Testing data

x → Learning f (f^{-1}) → ~~$P(y|x)$~~

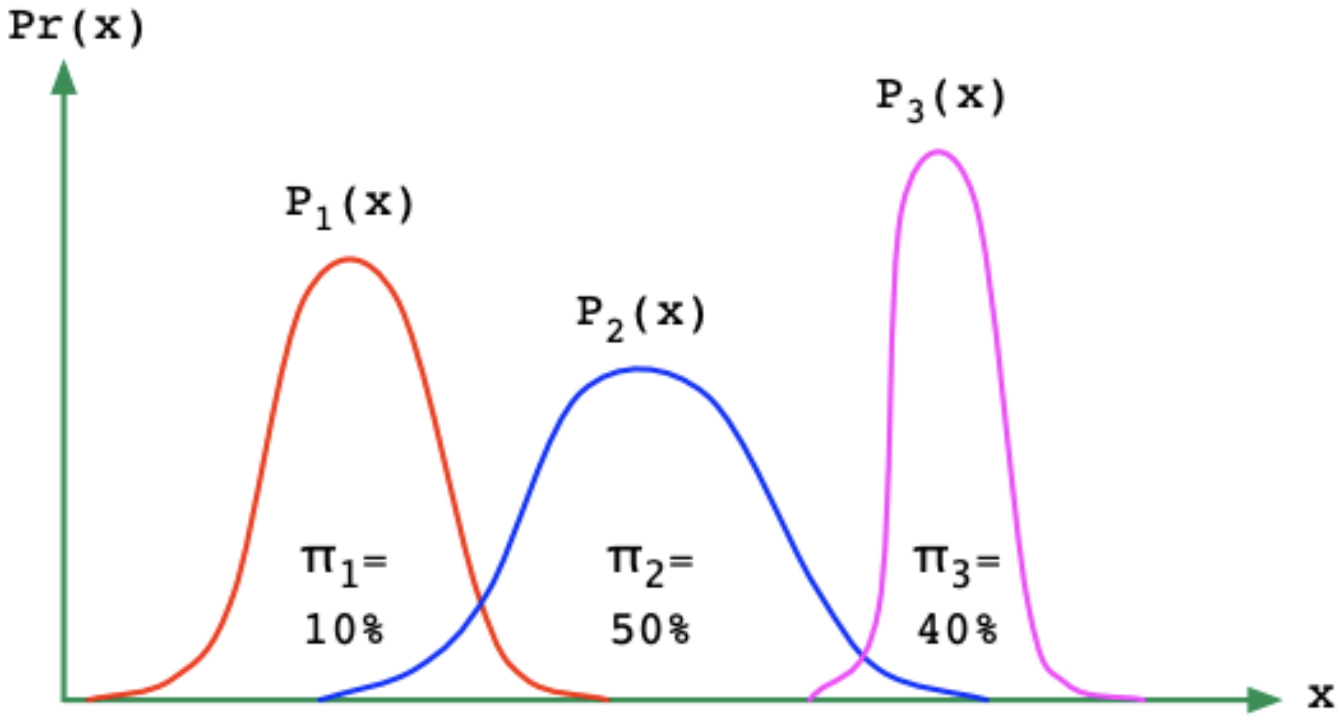
$P(x|y)$

Training data

Generative models

Q: how to classify a new data?

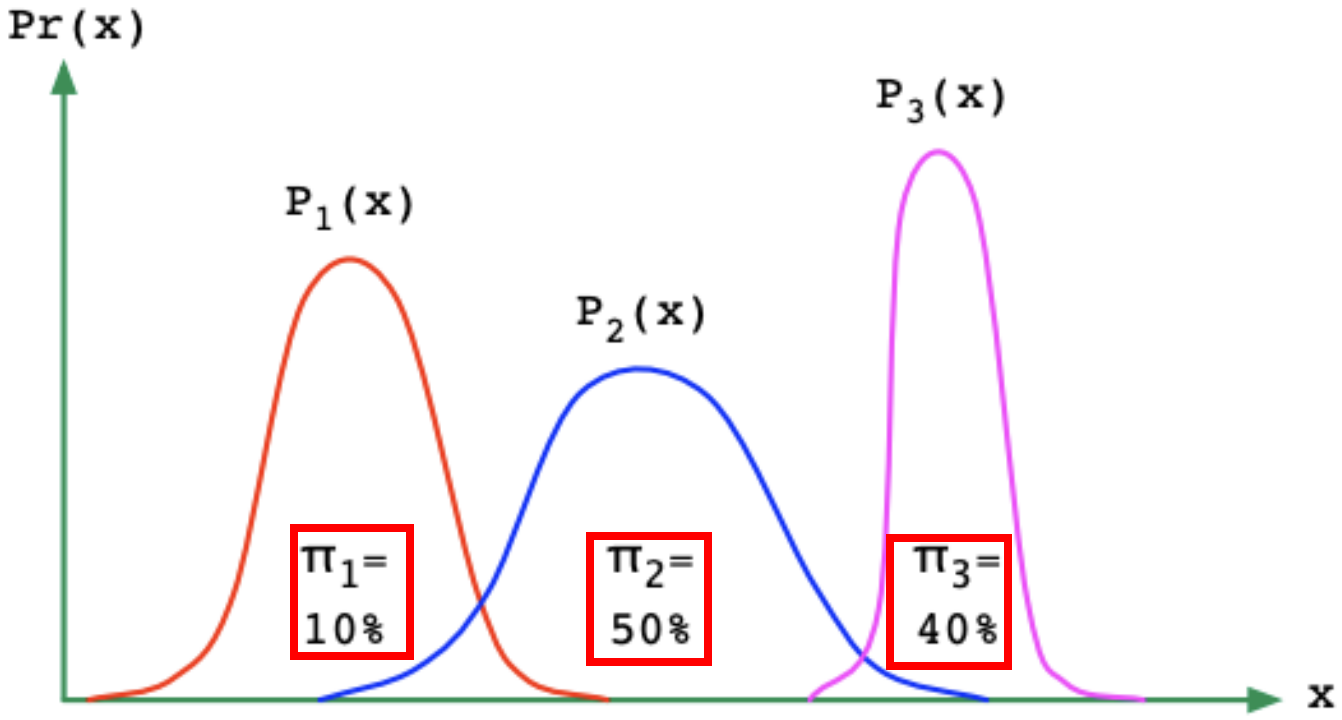
Generative models



$$\Pr(x, y) = \Pr(y)\Pr(x|y) = \pi_y P_y(x).$$

Image from https://cseweb.ucsd.edu/~dasgupta/250B/generative-intro-handout.pdf

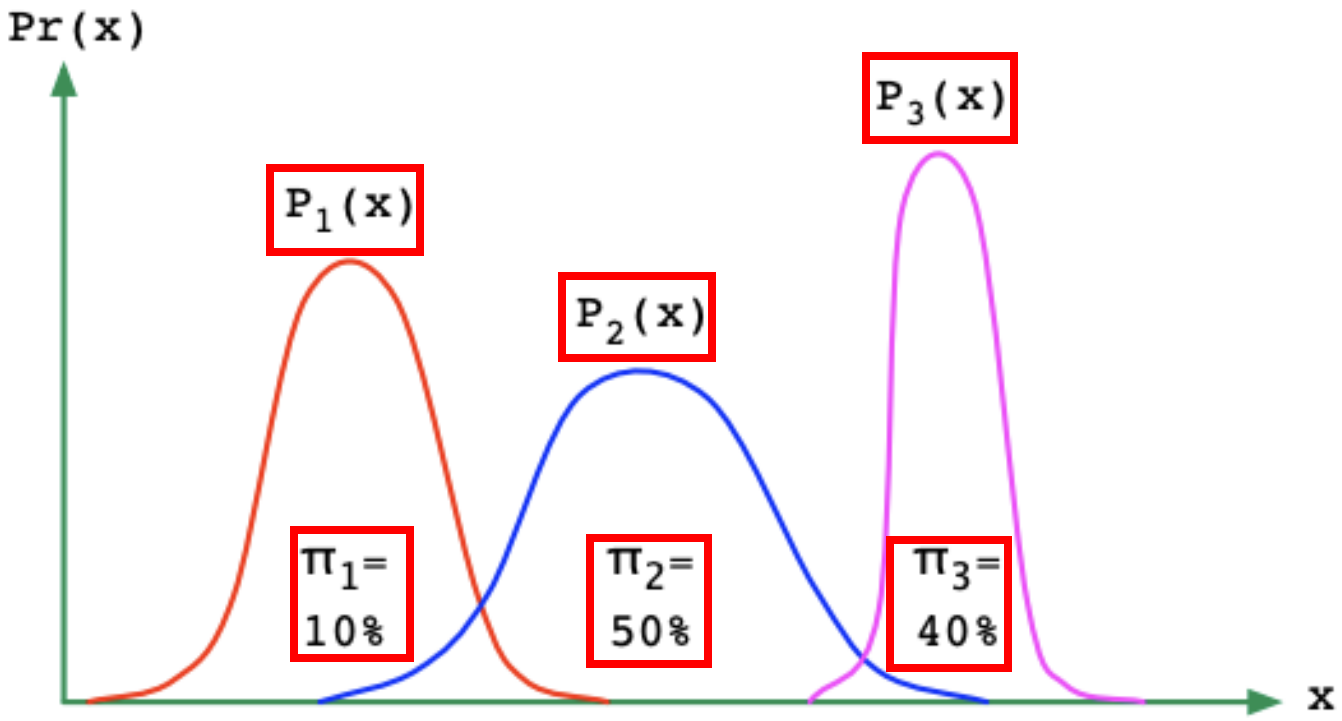
Generative models



$$\Pr(x, y) = \Pr(y)\Pr(x|y) = \pi_y P_y(x).$$

Image from https://cseweb.ucsd.edu/~dasgupta/250B/generative-intro-handout.pdf

Generative models

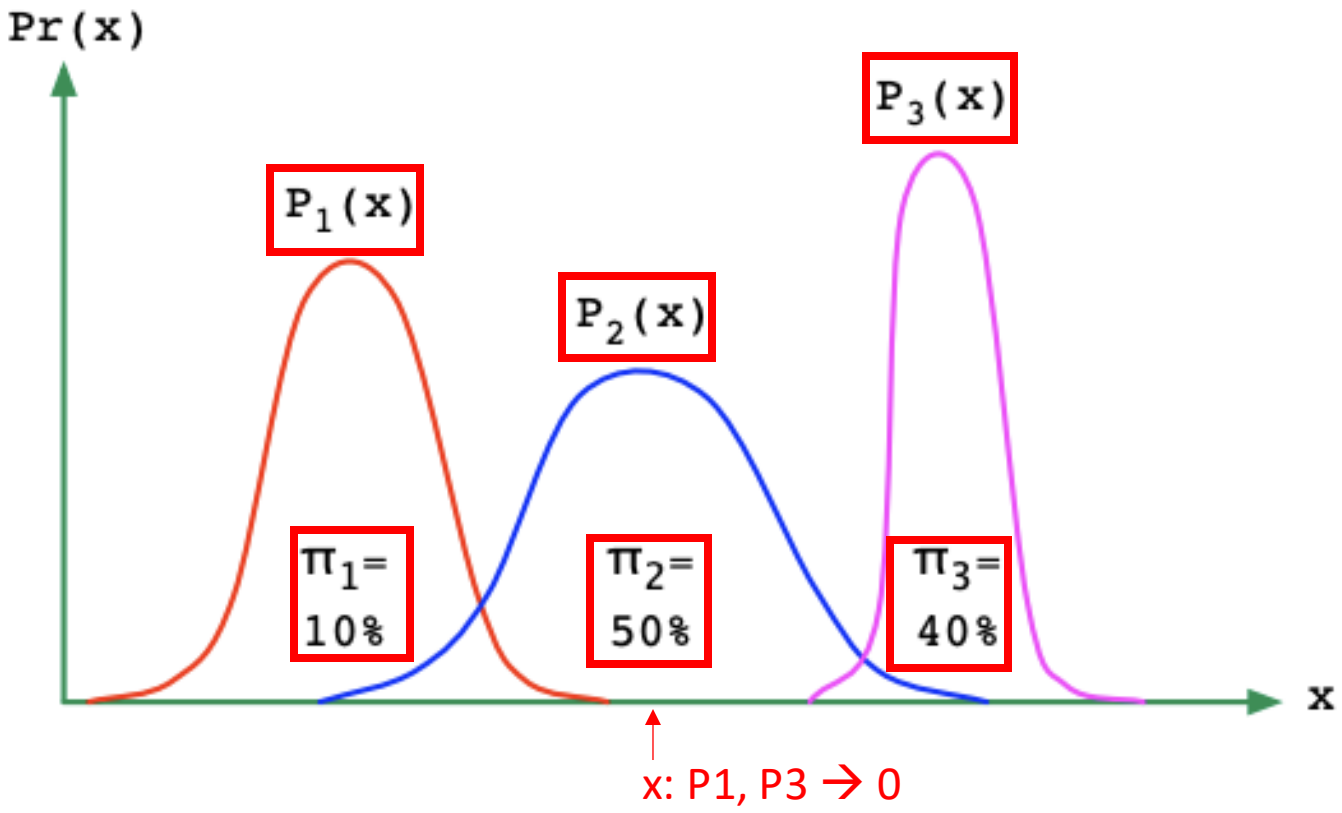


$$\Pr(x, y) = \Pr(y)\Pr(x|y) = \pi_y P_y(x).$$

Pick the label y making $P(x, y)$ largest

Image from https://cseweb.ucsd.edu/~dasgupta/250B/generative-intro-handout.pdf

Generative models

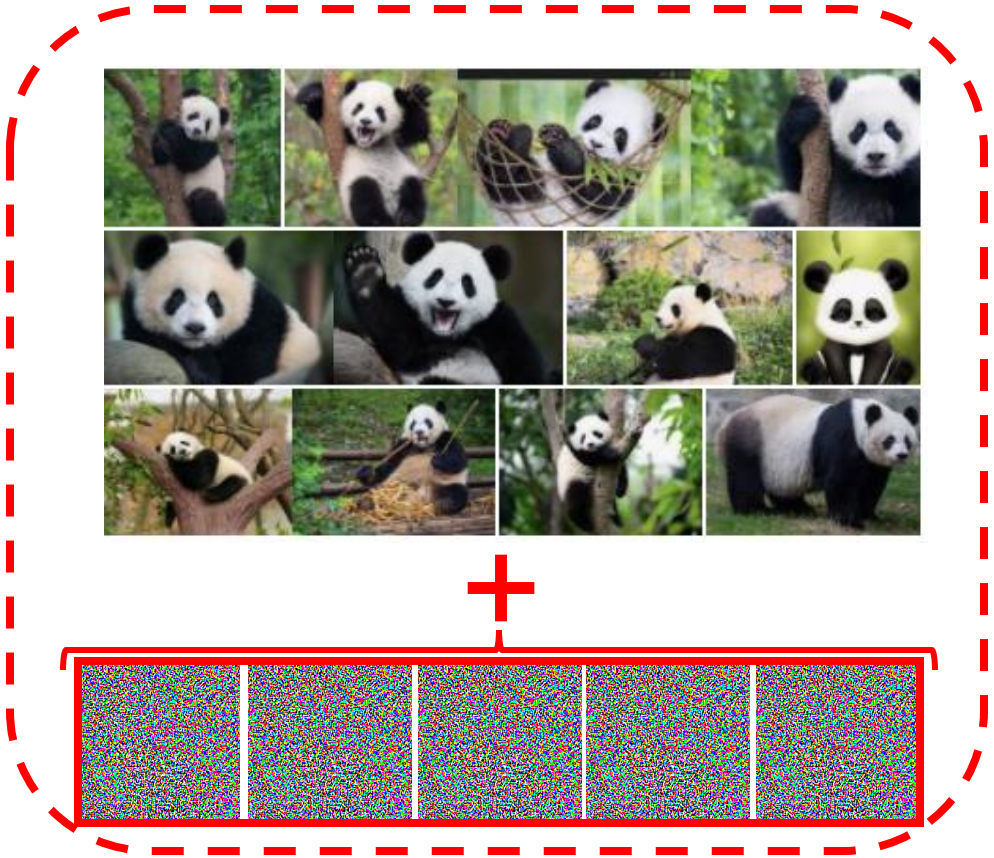


$$\Pr(x, y) = \Pr(y)\Pr(x|y) = \pi_y P_y(x).$$

Pick the label y making $P(x, y)$ largest

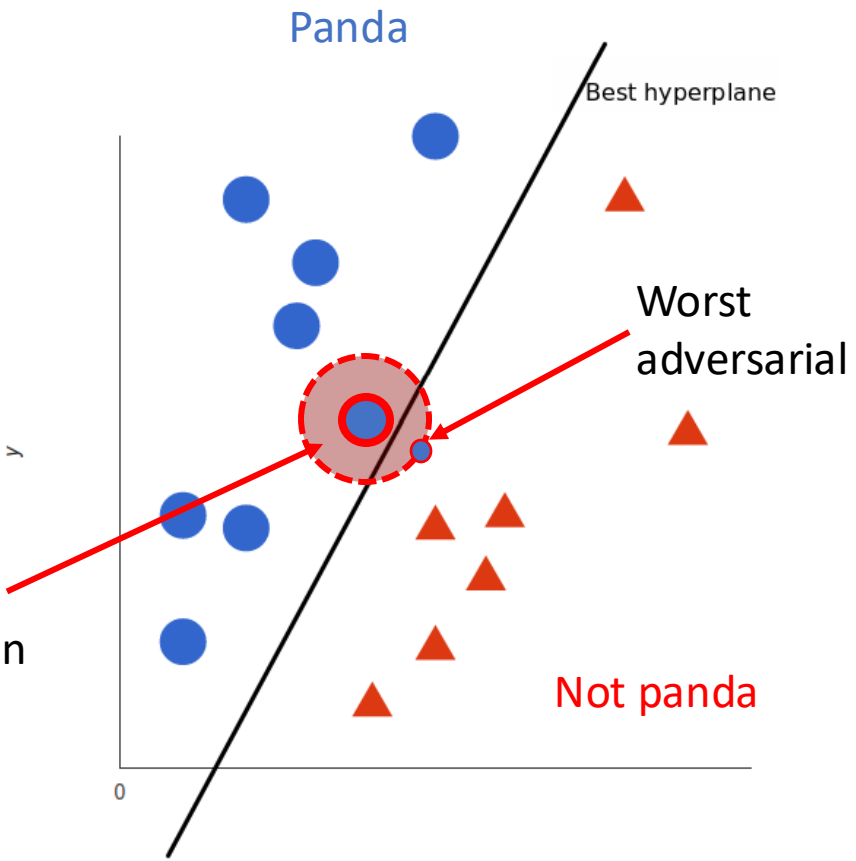
Image from https://cseweb.ucsd.edu/~dasgupta/250B/generative-intro-handout.pdf

Adversarial learning and generative models



Small modification (one pixel)

ML model for panda recognition



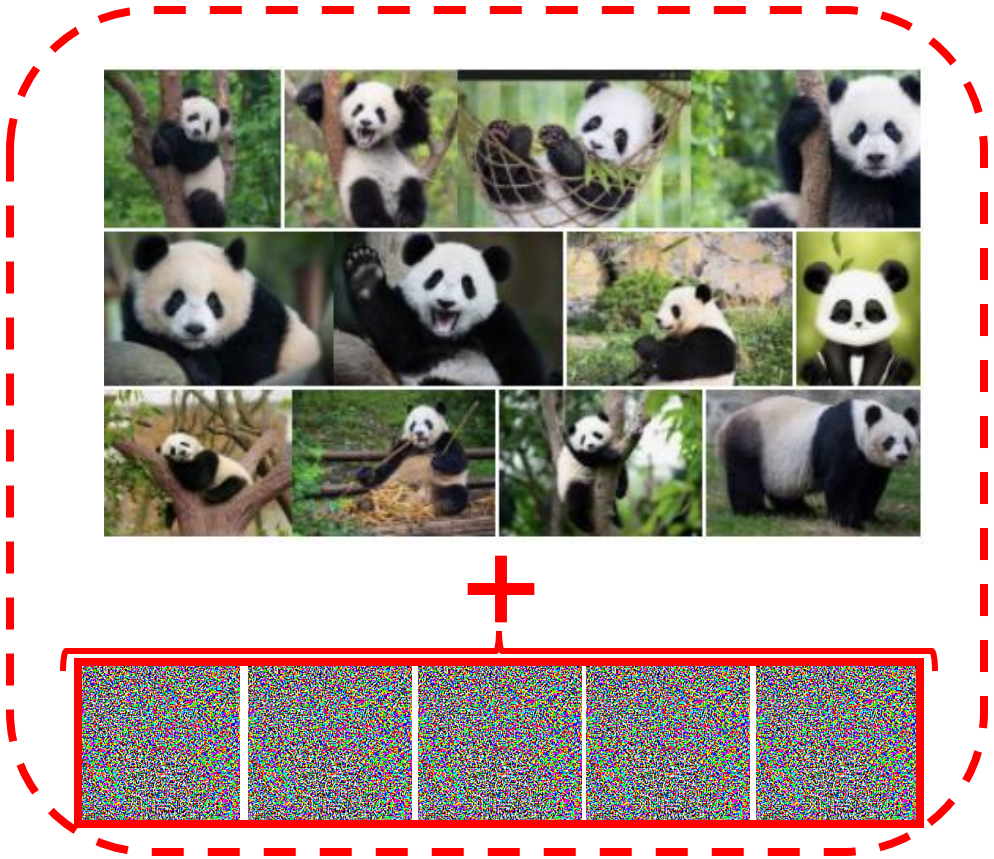
0~255

Q: for a grey scale images of size 28x28, how many possible one-pixel changes can we have?

255 x 28 x 28 = 199920

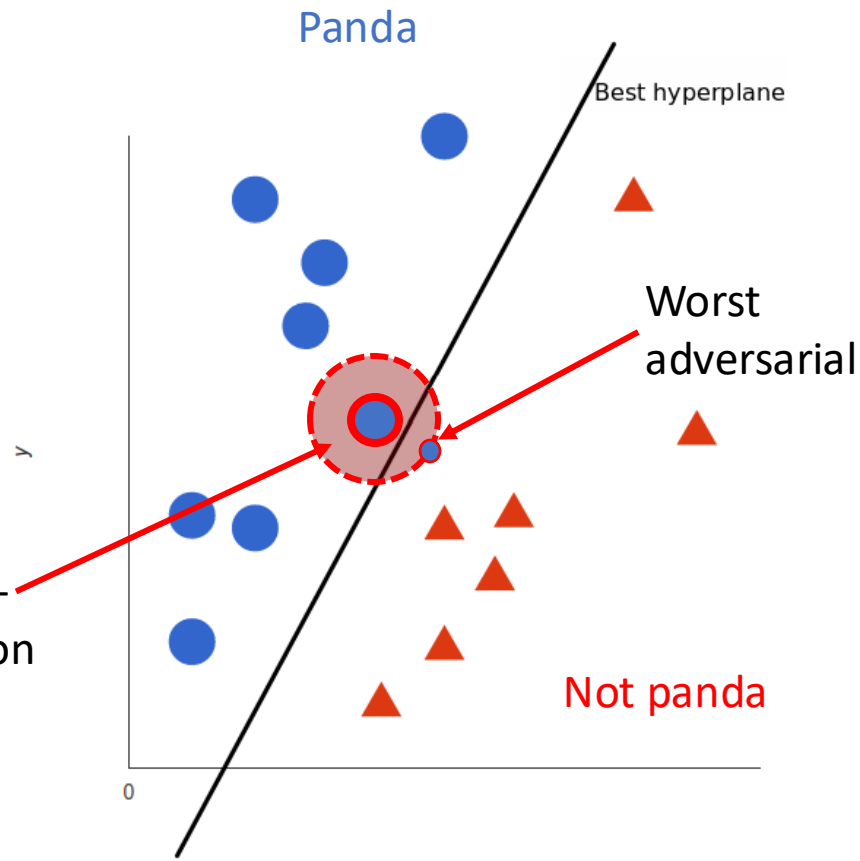
Adversarial learning and generative models

Limit: additive noise is only one way to generate adversarial data



Small modification (one pixel)

ML model for panda recognition



0~255

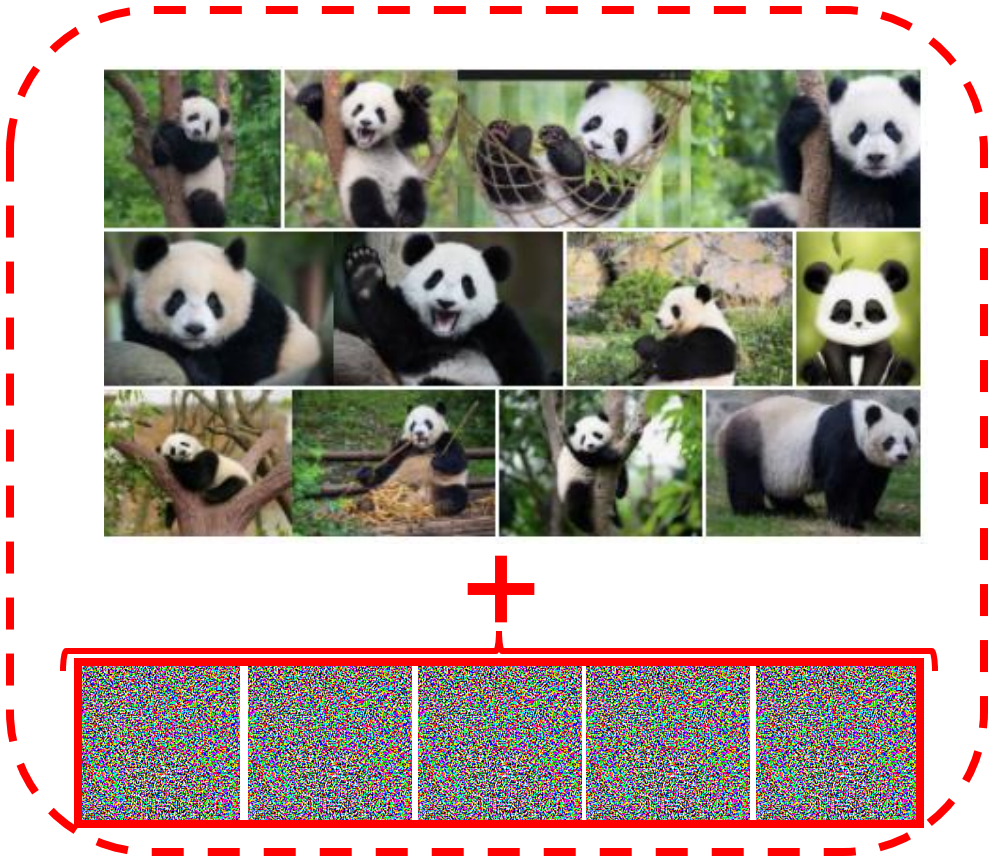
Q: for a grey scale images of size 28x28, how many possible one-pixel changes can we have?

255 x 28 x 28 = 199920

Adversarial learning and generative models

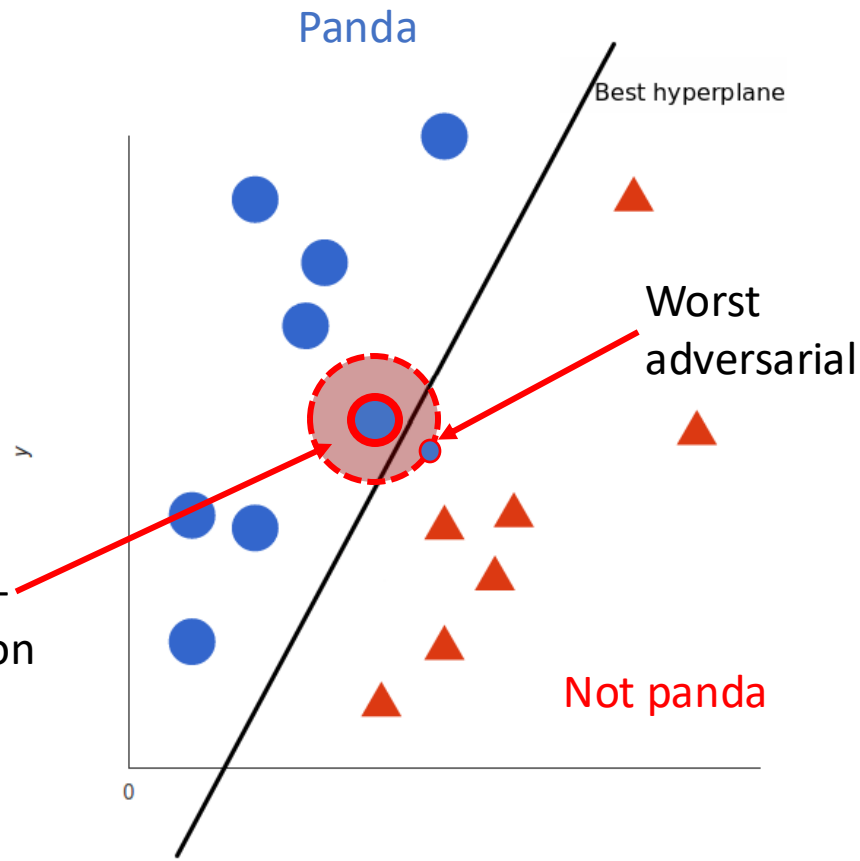
Limit: additive noise is only one way to generate adversarial data

Q: can generative models help?



Small modification (one pixel)

ML model for panda recognition



0~255

Q: for a grey scale images of size 28x28, how many possible one-pixel changes can we have?

255 x 28 x 28 = 199920

Generative adversarial networks

Network:
discriminative model

More details at tutorial note https://arxiv.org/pdf/1701.00160.pdf and original paper https://arxiv.org/pdf/1406.2661.pdf

Generative adversarial networks

Network:
discriminative model

Generative model:
generate data

More details at tutorial note https://arxiv.org/pdf/1701.00160.pdf and original paper https://arxiv.org/pdf/1406.2661.pdf

Generative adversarial networks

Target: better accuracy

Network:
discriminative model

Generative model:
generate data

More details at tutorial note https://arxiv.org/pdf/1701.00160.pdf and original paper https://arxiv.org/pdf/1406.2661.pdf

Generative adversarial networks

Target: better accuracy

Network:
discriminative model

Generative model:
generate data

Target of adversarial data:
worse accuracy

More details at tutorial note https://arxiv.org/pdf/1701.00160.pdf and original paper https://arxiv.org/pdf/1406.2661.pdf

Generative adversarial networks

Target: better accuracy

Network:
discriminative model

Generative model:
generate data

Target of adversarial data:
worse accuracy

$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

More details at tutorial note https://arxiv.org/pdf/1701.00160.pdf and original paper https://arxiv.org/pdf/1406.2661.pdf

Generative adversarial networks

Target: better accuracy

Network:
discriminative model

Generative model:
generate data

Target of adversarial data:
worse accuracy

$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

More details at tutorial note https://arxiv.org/pdf/1701.00160.pdf and original paper https://arxiv.org/pdf/1406.2661.pdf

Generative adversarial networks

Target: better accuracy

Network:
discriminative model

Generative model:
generate data

Target of adversarial data:
worse accuracy

$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

More details at tutorial note https://arxiv.org/pdf/1701.00160.pdf and original paper https://arxiv.org/pdf/1406.2661.pdf

Generative adversarial networks

Target: better accuracy

Network:
discriminative model

Generative model:
generate data

Target of adversarial data:
worse accuracy

$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

Use original training data

More details at tutorial note https://arxiv.org/pdf/1701.00160.pdf and original paper https://arxiv.org/pdf/1406.2661.pdf

Generative adversarial networks

Target: better accuracy

Network:
discriminative model

Generative model:
generate data

Target of adversarial data:
worse accuracy

$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

Use original training data

Maximum likelihood estimation

More details at tutorial note https://arxiv.org/pdf/1701.00160.pdf and original paper https://arxiv.org/pdf/1406.2661.pdf

Generative adversarial networks

Target: better accuracy

Network:
discriminative model

Generative model:
generate data

Target of adversarial data:
worse accuracy

$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

Use original training data
Maximum likelihood estimation

Input noise variable

More details at tutorial note https://arxiv.org/pdf/1701.00160.pdf and original paper https://arxiv.org/pdf/1406.2661.pdf

Generative adversarial networks

Target: better accuracy

Network:
discriminative model

Generative model:
generate data

Target of adversarial data:
worse accuracy

$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

Use original training data
Maximum likelihood estimation

Input noise variable
Use generated data (from noise)

More details at tutorial note https://arxiv.org/pdf/1701.00160.pdf and original paper https://arxiv.org/pdf/1406.2661.pdf

Generative adversarial networks

Target: better accuracy

Network:
discriminative model

Generative model:
generate data

Target of adversarial data:
worse accuracy

$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

Use original training data
Maximum likelihood estimation

Input noise variable
Use generated data (from noise)
Generated data → discriminator

More details at tutorial note https://arxiv.org/pdf/1701.00160.pdf and original paper https://arxiv.org/pdf/1406.2661.pdf

Generative adversarial networks

Target: better accuracy

Network:
discriminative model

Generative model:
generate data

Target of adversarial data:
worse accuracy

$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))]$$

MLE of –D(adv.)

Use original training data
Maximum likelihood estimation

Input noise variable
Use generated data (from noise)
Generated data → discriminator

More details at tutorial note https://arxiv.org/pdf/1701.00160.pdf and original paper https://arxiv.org/pdf/1406.2661.pdf

Generative adversarial networks

Target: better accuracy

Network:
discriminative model

Generative model:
generate data

Target of adversarial data:
worse accuracy

$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))]$$

MLE of –D(adv.)

Use original training data
Maximum likelihood estimation

Input noise variable
Use generated data (from noise)
Generated data → discriminator

More details at tutorial note https://arxiv.org/pdf/1701.00160.pdf and original paper https://arxiv.org/pdf/1406.2661.pdf