EURO^2

Fundamental Concepts of Generative Machine Learning

Erdem Akagündüz, PhD
Graduate School of Informatics, METU, Türkiye
ncc@ulakbim.gov.tr

Lesson 1: Mathematical Background

Welcome to Part I: “Mathematical Background”

This part includes four subsections:

- **Generation vs. Discrimination in Machine Learning**
- Data Distribution, Sampling, Inference and Generation
- Expectation and Likelihood
- Evaluation for Generative Models, Distribution Distances, Divergence and Entropy

Discriminative vs Generative Models

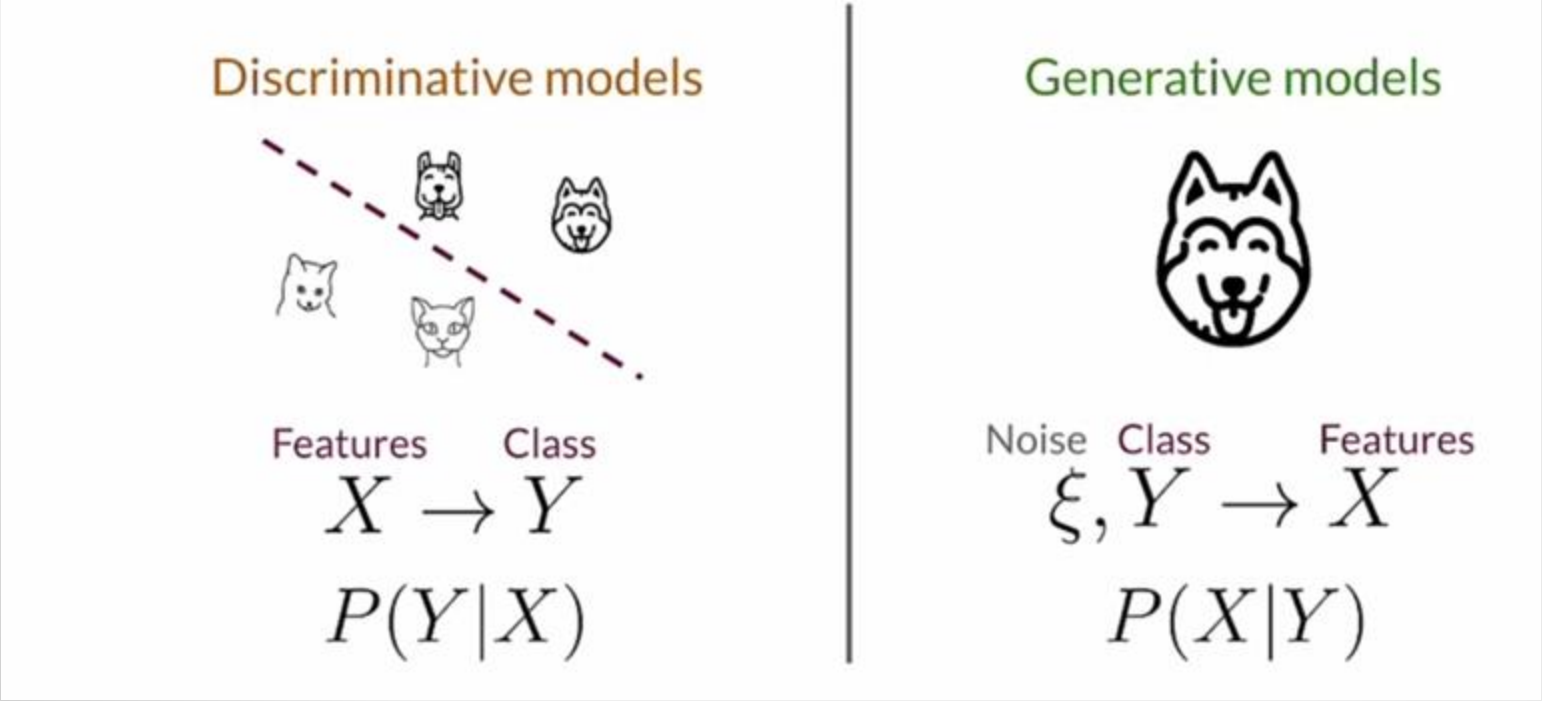
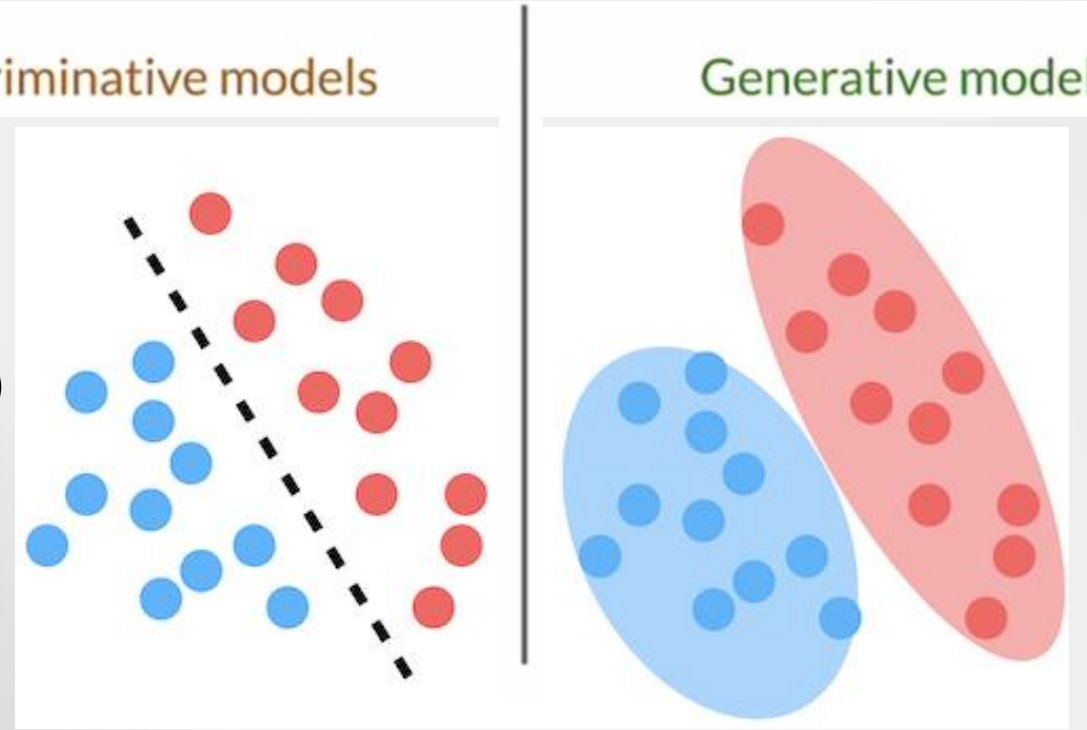


figure by coursera.org

Discriminative vs Generative Models

- Logistic regression
- Scalar Vector Machine
- Neural networks (CNNs)
- etc



- Naïve Bayes
- Bayesian networks
- Markov random fields
- Hidden Markov Models
- Gaussian Mixture Models
- etc

“decision boundary”

“distribution”

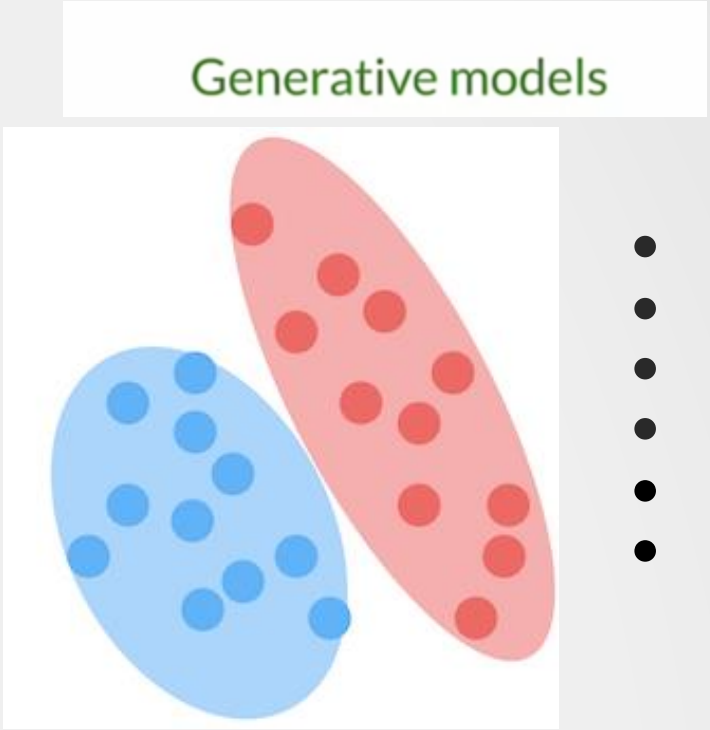
Before the age of GANs, conventional ML scrutinised classifiers into these two categories.

Generative Classifier

A “generative classifier” is a statistical model of the joint probability distribution:

$$P(X,Y)$$

- on given observable variable $\mathbf{X}$
  - Observation $\stackrel{\text{def}}{=}$ Feature
- and target variable $\mathbf{Y}$



“distribution”

- Naïve Bayes
- Bayesian networks
- Markov random fields
- Hidden Markov Models
- Gaussian Mixture Models
- etc

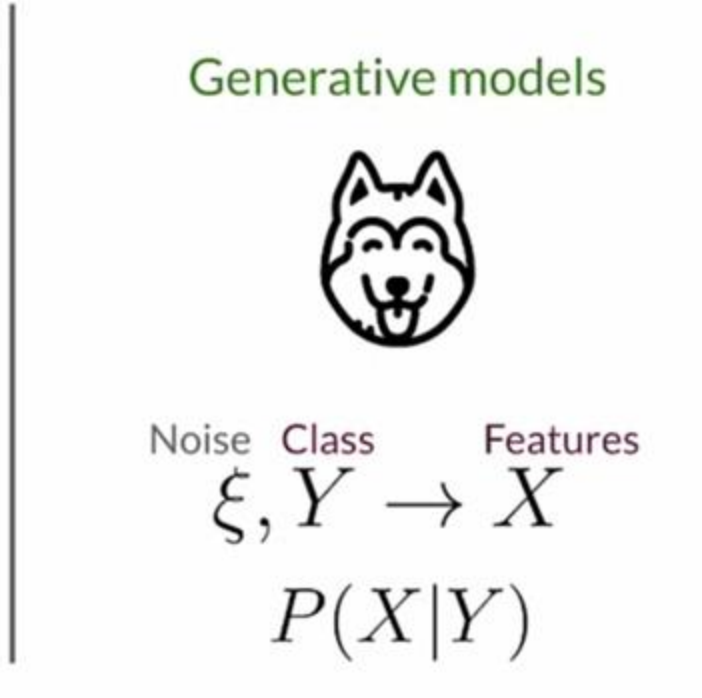
Generative Classifier (vs Generation)

A “generative classifier” is a statistical model of the joint probability distribution:

$$P(X,Y)$$

- on given observable variable $\mathbf{X}$
  - Observation $\stackrel{\text{def}}{=}$ Feature
- and target variable $\mathbf{Y}$

“*When a new observation is fed to a generative classifier, it tries to predict which class would have most likely generated the given observation.*"



“distribution”

- Naïve Bayes
- Bayesian networks
- Markov random fields
- Hidden Markov Models
- Gaussian Mixture Models
- etc

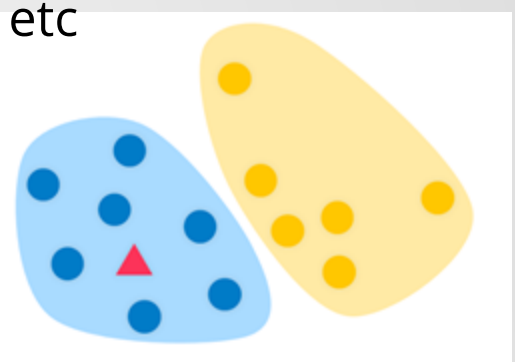
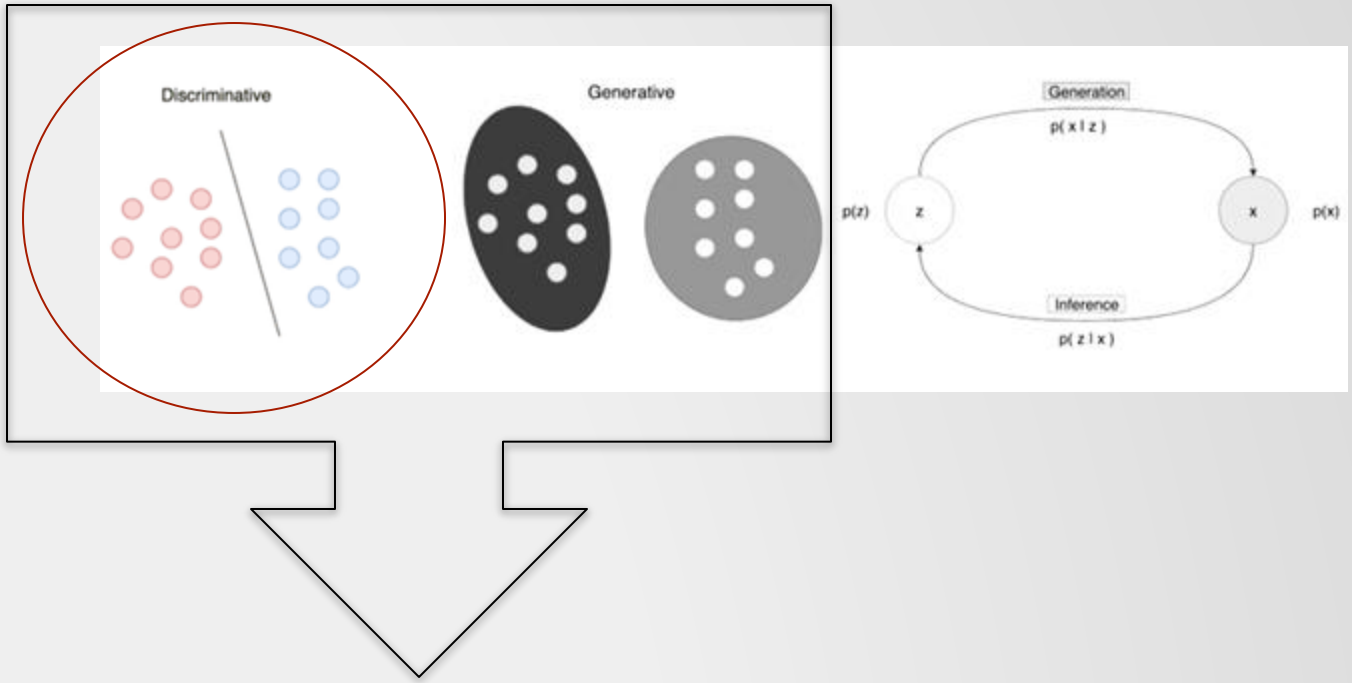


figure by coursera.org

Classification vs Generation

Discriminative models (or classifiers) are machine learning models that learn to classify input data into different categories based on a set of features.



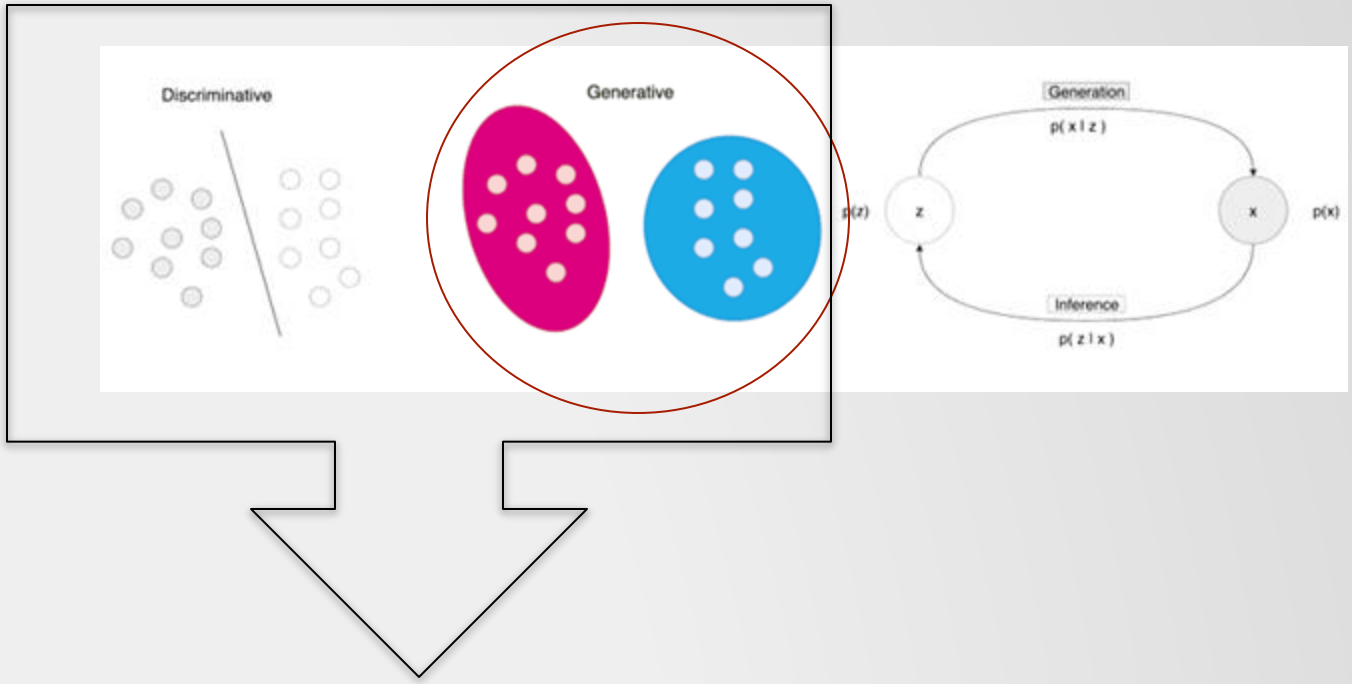
classifiers

These models learn to distinguish between different classes of data **without necessarily modeling the underlying probability distribution of the data**.

figure by coursera.org

Classification vs Generation

Generative classifiers are models that **learn to model the probability distribution of the input data for each class separately, and then use Bayes' rule to calculate the posterior probability of each class given the input**.



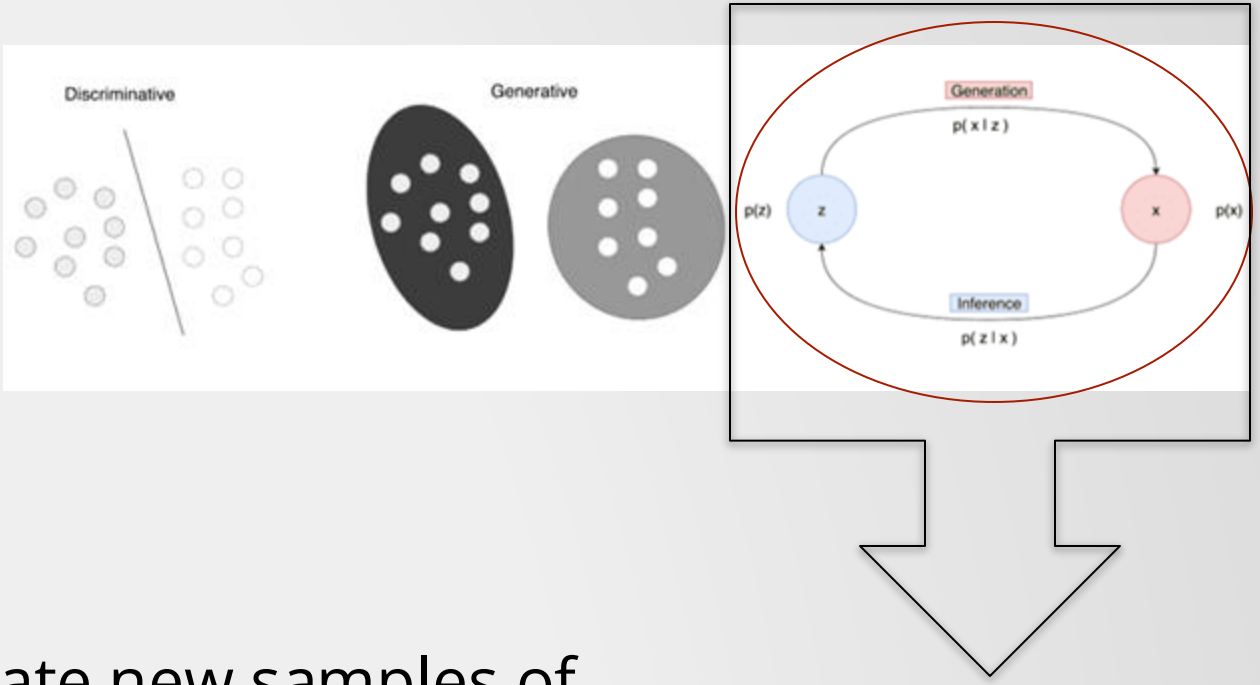
classifiers

These models can be used to classify new data points into different categories based on their likelihood under each class distribution.

figure by coursera.org

Classification vs Generation

Generative (or generation) models are **machine learning models that learn to model the probability distribution of the input data.**



These models can be used to generate new samples of data that resemble the original data distribution.

generators

figure by coursera.org

Generation

- $p(z)$: a noise distribution (i.e. latent space)
- $p(x|z)$: is the generation process, i.e. the generation of a new sample x, given a noise vector z
- $p(x)$: is the data/observation distribution.
- $p(z|x)$: is the inference process, i.e. extracting the noise vector z that would generate the sample x.

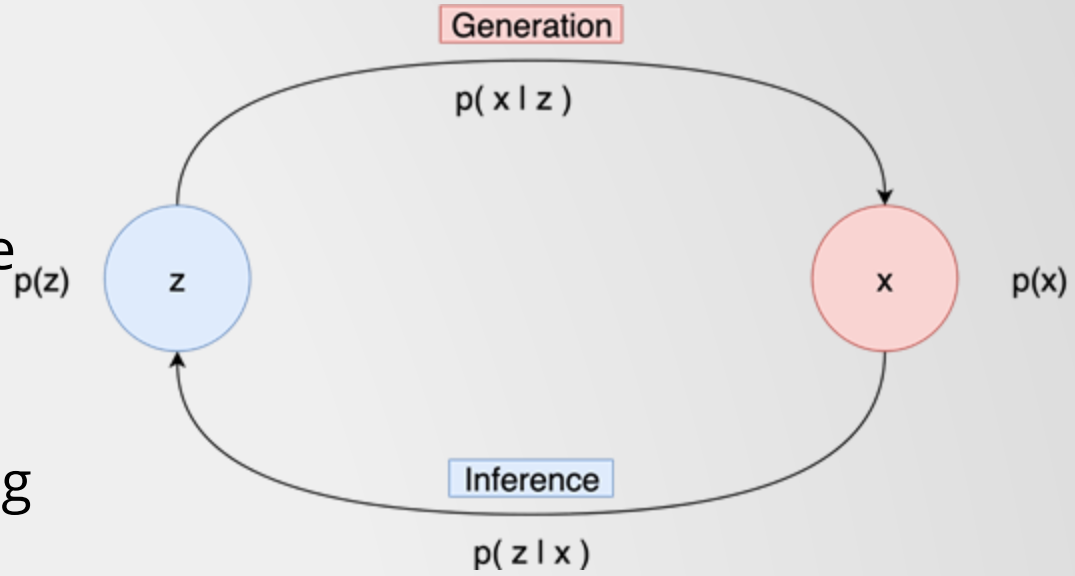


figure by coursera.org

Next lecture:

Part I: “Mathematical Background”

- Generation vs. Discrimination in Machine Learning
- **Data Distribution, Sampling, Inference and Generation**
- Expectation and Likelihood
- Evaluation for Generative Models, Distribution Distances, Divergence and Entropy

Thanks

This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 101101903. The JU receives support from the Digital Europe Programme and Germany, Bulgaria, Austria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Latvia, Poland, Portugal, Romania, Slovenia, Spain, Sweden, France, Netherlands, Belgium, Luxembourg, Slovakia, Norway, Türkiye, Republic of North Macedonia, Iceland, Montenegro, Serbia