TÜBİTAK
ODTÜ METU
INFORMATICS INSTITUTE

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

The theorem is named after Reverend Thomas Bayes, an 18th-century British statistician and theologian.

Bayes Theorem

- Bayes' Theorem is a fundamental concept in probability theory that describes the probability of an event based on prior knowledge of related events.
- It provides a way to update our beliefs about the probability of an event as new evidence is obtained.

Independent Events

$$P(X \cap Y) = P(X) \bullet P(Y)$$

Dependent Events

$$P(X \cap Y) = P(Y) \bullet P(X \mid Y)$$

LIKELIHOOD
The probability of "B" being True, given "A" is True

PRIOR
The probability "A" being True. This is the knowledge.

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)}$$

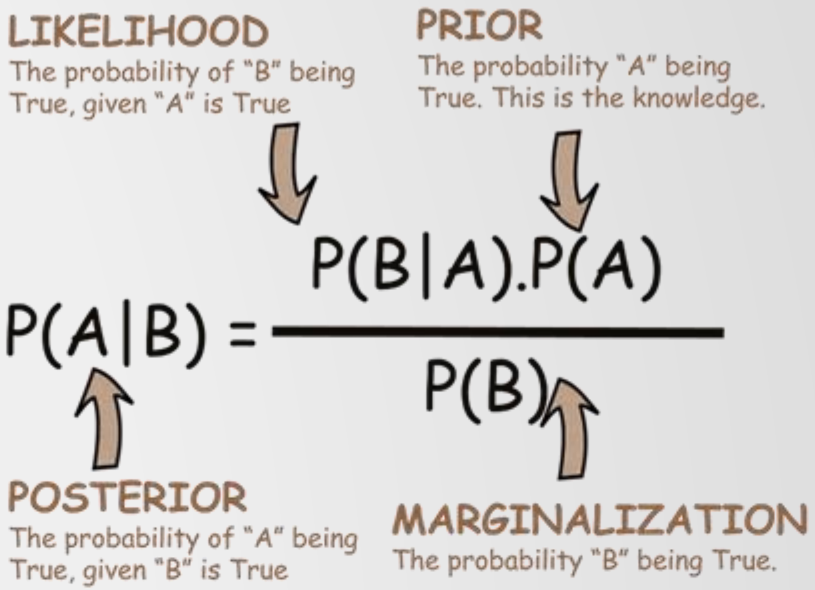
POSTERIOR
The probability of "A" being True, given "B" is True

MARGINALIZATION
The probability "B" being True.
also called the "Evidence"

The theorem is named after Reverend Thomas Bayes, an 18th-century British statistician and theologian.

Bayes Theorem

- Bayes Theorem plays a crucial role in generative models, which are used to learn the underlying probability distributions of a given dataset.
- Bayes Theorem provides a way to update our prior beliefs about the parameters of a probability distribution in light of new evidence (i.e., data).
- In the context of generative models, the theorem is used to estimate the parameters of the underlying distribution that generated the data.
- Specifically, it helps us to update our prior beliefs about the parameters of the distribution based on the observed data.

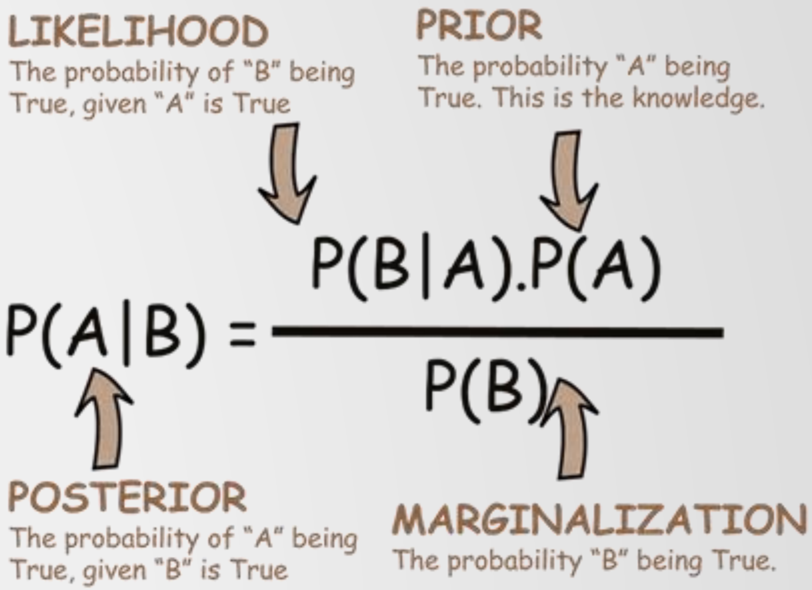


also called the "**Data**"

The theorem is named after Reverend Thomas Bayes, an 18th-century British statistician and theologian.

Bayes Theorem

- Bayes Theorem plays a crucial role in generative models, which are used to learn the underlying probability distributions of a given dataset.
- Bayes' Theorem is widely used in deep learning-based generative models, such as GANs, VAEs, and Bayesian neural networks.
- In GANs, for example, the discriminator network can be seen as an approximate likelihood function, and the generator network is used to generate samples from the learned posterior distribution over the latent variables.

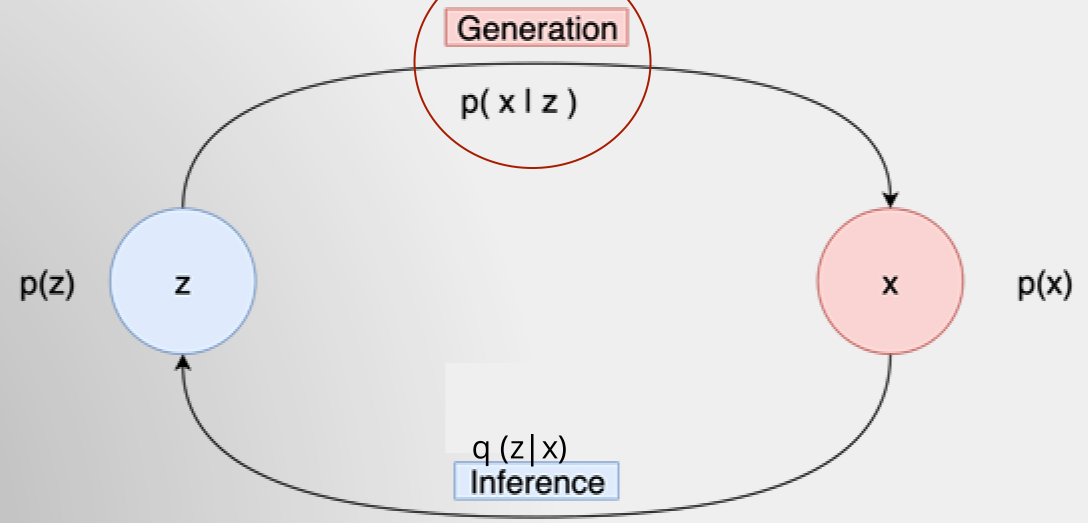


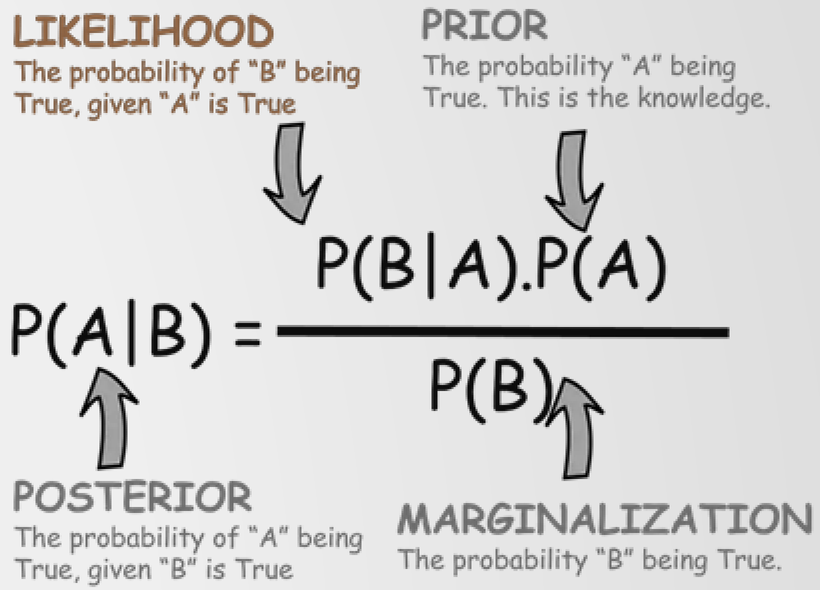
also called the "**Data**"

Likelihood?

In the context of generative models, the likelihood refers to **the probability of observing a given set of data points under the assumed probability distribution of the generative model.**

In other words, the likelihood measures how well the generative model can explain the observed data.





Maximizing the Likelihood

- Maximizing the likelihood involves finding the set of model parameters θ that maximize the likelihood function (i.e. the Generator)

  For example, by using the log-likelihood loss function in a deep generative model to maximize the likelihood of the model parameters given the data.

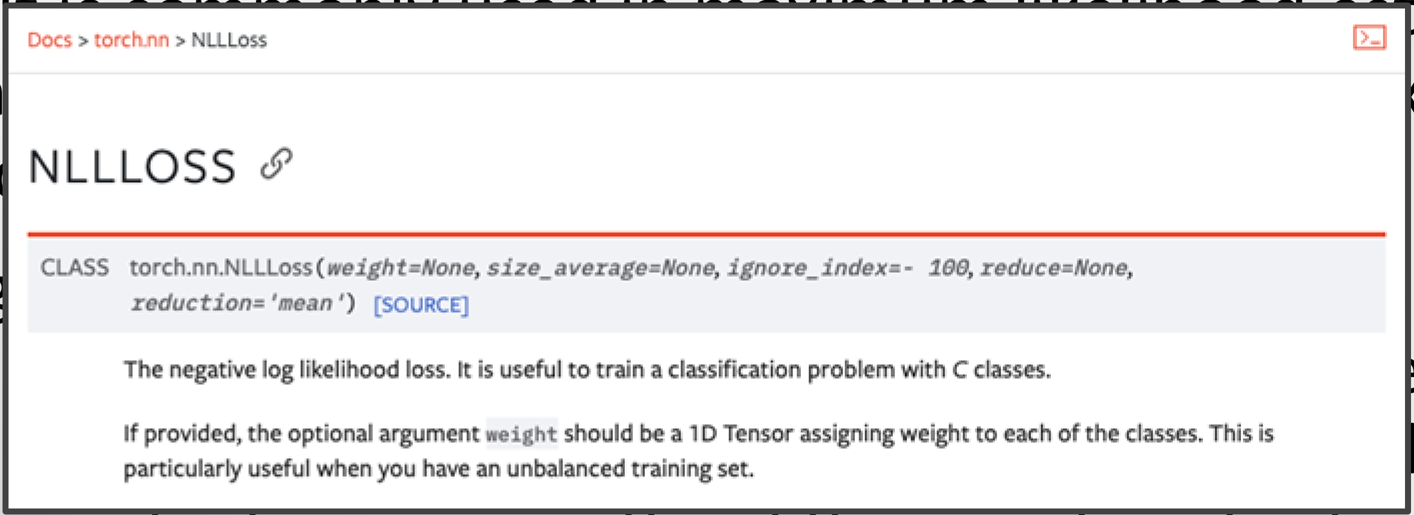
- In deep generative models, the likelihood function is often intractable or difficult to optimize directly.

  For example, in VAEs, a lower bound on the likelihood is optimized instead, while in GANs, a game-theoretic objective is used to implicitly maximize the likelihood.

- Maximizing likelihood is a common objective in deep generative models.

Log-likelihood Loss function

- Log-likelihood is a measure of how well a statistical model fits the data it is given. … imation (MLE), where th … imize the likelihoo …


Docs > torch.nn > NLLLoss

NLLLOSS

CLASS torch.nn.NLLLoss(*weight=None, size_average=None, ignore_index=- 100, reduce=None, reduction='mean'*) [SOURCE]

The negative log likelihood loss. It is useful to train a classification problem with C classes.

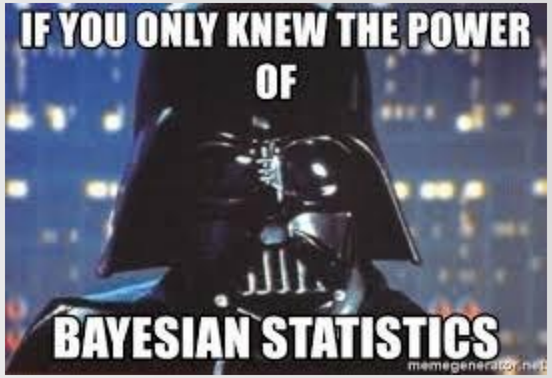
If provided, the optional argument `weight` should be a 1D Tensor assigning weight to each of the classes. This is particularly useful when you have an unbalanced training set.

- In deep g … a loss function to train t … difference between … istribution of the data. The goal is to minimize this difference by adjusting the model's parameters.

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