# Energy-Based Models in Self-Supervised Learning

Energy-Based Models (EBMs) provide a flexible framework for defining probability distributions in self-supervised learning tasks. The key idea is to define an **energy function** $E(x, y)$ that measures the compatibility between observed data $x$ and the target $y$ to be predicted. Some key properties of EBMs include:

- The energy function outputs a scalar value, with lower values indicating higher compatibility between $x$ and $y$.
- The probability of a target $y$ given data $x$ is defined as:

$$p(y|x) = \frac{\exp(-E(x, y))}{\int_{y'} \exp(-E(x, y'))dy'} \tag{1}$$

  where the denominator is an intractable normalization constant.

- Training an EBM involves minimizing the energy of observed data pairs $(x, y)$ while maximizing the energy for incompatible pairs.
- EBMs avoid the need to explicitly model the full joint distribution $p(x, y)$, allowing more flexibility in model design.

EBMs provide a unifying framework for many self-supervised learning methods:

- **Contrastive methods** push down the energy of observed data pairs and push up the energy of carefully selected negative pairs.
- **Denoising autoencoders** learn an energy function that assigns low energy to clean data and higher energy to corrupted versions.
- **Generative Adversarial Networks (GANs)** can be interpreted as learning an energy function (the discriminator) that assigns low energy to real data and higher energy to generated samples.

In summary, EBMs provide a powerful and flexible framework for defining probability distributions in self-supervised learning. By learning an energy function that assigns low values to compatible data pairs, EBMs can learn useful representations from unlabeled data without the need for explicit probabilistic modeling.