IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

EXPLORING ENERGY BASED MODELS FOR GENERATIVE MODELLING

[1]Sushruth S, [2]Mohammed Mutahar,[3]Vallepali Jatin Chandra, [4]Sanjiv MV, [5]Deepthi S

[1,2,3,4]B.Tech CSE Presidency University, [5]Assistant Professor Dept of CSE Presidency University Bengaluru, India

Abstract: Energy-based models (EBMs) offer a promising approach to generative modeling by leveraging Boltzmann distribution to express the probability of events. In this study, we delve into the theory and practical implementation of EBMs, drawing inspiration from physical systems and neural network architectures. By training neural networks to output low scores for likely observations and high scores for unlikely ones, EBMs aim to model the true data-generating distribution. We address challenges associated with sampling new observations and the intractable normalizing denominator, proposing approximation techniques like contrastive divergence and Langevin dynamics. Through exploration and implementation, we aim to provide insights into the construction and utilization of EBMs for image generation tasks.

Index Terms – Energy-based models, Boltzmann distribution, neural network, contrastive divergence, Langevin dynamics.

I. Introduction

Energy-based models (EBMs) offer a novel paradigm in generative modeling, inspired by Ludwig Boltzmann's distribution formulation for describing gasses in thermal equilibrium. By employing neural networks to define energy functions, EBMs strive to model the true data-generating distribution. However, challenges arise concerning sampling new observations and computing the intractable normalizing denominator. To overcome these hurdles, we delve into approximation techniques such as contrastive divergence and Langevin dynamics. Through an illustrative narrative and theoretical exploration, this study aims to elucidate the fundamental concepts behind EBMs and their practical implementation for generating images of handwritten digits.

II.Objective

This study aims to provide a comprehensive understanding of energy-based models (EBMs) and their application to generative modeling tasks, particularly in the context of image generation. By exploring the theory and practical implementation of EBMs, we seek to address challenges related to sampling new observations and computing the intractable normalizing denominator. Through experimentation and code implementation, our objective is to equip readers with the knowledge and tools necessary to construct and utilize EBMs effectively for image generation purposes.

III.Literature Review

“Learning Energy-Based Models in High-Dimensional Spaces with Multiscale Denoising-Score Matching" by Zengyi Li et al. delves into the potential of energy-based models (EBMs) across diverse applications, including sample synthesis and denoising tasks. Conventional training methods for EBMs suffer from sluggishness due to extensive sampling requirements. Denoising-score matching emerges as a swifter alternative, yet its effectiveness is hampered in high-dimensional data settings by single noise-level training. Addressing this challenge, the paper introduces multiscale denoising-score matching (MDSM) to enhance sample synthesis quality, rivaling state-of-the-art techniques such as GANs. Theoretical insights into score matching, particularly denoising-score matching, are thoroughly explored. Techniques like double backpropagation are employed to optimize objectives involving neural networks. Prior research underscores the imperative of training with multiple noise levels for optimal EBM performance. Building upon this foundation, the study proposes MDSM and validates its efficacy through comprehensive empirical evaluations.

"Your Classifier Is Secretly An Energy Based Model And You Should Treat It Like One" by Will Grathwohl et al. delves into the reinterpretation of a standard discriminative classifier as an energy-based model for the joint distribution of data and labels, aiming to enhance various aspects of model performance. By adopting energy-based training, the model exhibits improvements in calibration, robustness, and out-of-distribution detection while generating samples comparable to recent GAN approaches. This framework enables training on unlabeled data and utilizes standard classifier architectures, offering a promising avenue for hybrid generative-discriminative models. Despite the inherent challenges of working with energy-based models, they fit more seamlessly within a discriminative framework and facilitate the use of modern classifier architectures. The contributions of this paper include a novel framework for joint modeling of labels and data, surpassing previous hybrid models in both generative and discriminative tasks, and showcasing improved calibration and robustness. Additionally, the paper discusses the theoretical foundations of energy-based models and recent advancements in training large-scale EBMs on high-dimensional data using techniques like Stochastic Gradient Langevin Dynamics.

"How to Train Your Energy-Based Models" by Yang Song et al. delves into modern approaches for training Energy-Based Models (EBMs), which offer flexibility in modeling probability distributions but pose challenges due to the unknown normalizing constant. It provides a beginner-friendly overview of training methods, starting with maximum likelihood training using Markov chain Monte Carlo (MCMC), then delving into MCMC-free approaches like Score Matching (SM) and Noise Contrastive Estimation (NCE). Theoretical connections among these methods are highlighted, and alternative training methods still under active research are briefly surveyed. Targeted at audiences with a basic understanding of generative models, the tutorial aims to facilitate the application of EBMs or initiate research projects in this domain. EBMs, less restrictive in functional form compared to models with tractable likelihoods, find applications across various fields including image generation, discriminative learning, natural language processing, density estimation, and reinforcement learning. Despite their modeling advantages, the exact likelihood computation and sample synthesis from EBMs remain challenging, necessitating innovative training approaches like MCMC, SM, and NCE, which are elucidated in this paper.

"Energy-Based Models for Sparse Overcomplete Representations" by Yee Whye Teh et al. delves into a novel extension of Independent Components Analysis (ICA) to overcomplete representations, departing from traditional causal generative approaches. Features are defined as deterministic functions of inputs, leading to marginal dependencies among features but conditional independence given inputs. By assigning energies to features, a probability distribution over input states is defined via the Boltzmann distribution, with parameters trained using contrastive divergence. Experimentally, the proposed learning algorithm demonstrates blind source separation on speech data and feature extraction from various datasets. The paper contrasts three main approaches to ICA: information maximization, causal generative, and energy-based models, proposing the latter as a tractable way to extend ICA to overcomplete and multi-layer models. The energy-based approach maintains computational convenience and proper density modeling in the overcomplete case, with features serving as deterministic functions of observation vectors. While the causal generative approach assumes marginal independence of sources, the energy-based approach prioritizes modeling probability density, making it suitable for applications requiring interpretable data structures or useful representations for action control.

"Energy-Based Models in Document Recognition and Computer Vision" by Yann Lecun et al. addresses two critical challenges in machine learning and pattern recognition: the normalization problem and the deep learning problem. It explores the development of "un-normalized" learning models, originally conceived in the handwriting recognition community during the 90s, which include graph transformer networks, conditional random fields, hidden Markov SVMs, and maximum margin Markov networks. These models are unified under the framework of energy-based models (EBM). Additionally, the paper discusses the deep learning problem, emphasizing the need for integrated training of all levels of recognition systems. Traditional methods like convolutional networks and back-propagation are effective but require extensive training data. The paper proposes leveraging unsupervised learning to initialize layers in deep networks, significantly reducing the need for training samples, especially in tasks like category-level recognition of everyday objects.

IV. METHODOLOGY

Energy-based models (EBMs) represent a paradigm shift in generative modeling, drawing inspiration from both physical systems and neural network architectures. At the core of EBMs lies the Boltzmann distribution, a fundamental concept in statistical physics, which offers a framework for expressing the probability of events based on their energy levels. This distribution provides a mathematical formulation that allows EBMs to assign probabilities to different configurations of data.

In EBMs, neural networks play a central role as they are trained to assign low energy (or scores) to likely observations and high energy to unlikely ones. By learning to discriminate between plausible and implausible data configurations, neural networks enable EBMs to capture the underlying data distribution effectively. This training process involves adjusting the parameters of the neural network to minimize the energy assigned to observed data points while maximizing the energy for unobserved or unlikely configurations.

However, training EBMs poses several challenges, the foremost being sampling new observations. Traditional methods for sampling from EBMs rely on approximation techniques such as Langevin dynamics. Langevin dynamics simulate a Markov chain to generate samples by iteratively updating the input data based on the gradient of the energy function. While effective, Langevin dynamics can be computationally expensive, especially for high-dimensional data.

Another challenge in training EBMs is the intractable normalizing denominator in the Boltzmann distribution. This term represents the sum or integral of the energies of all possible data configurations and is essential for ensuring that the probabilities computed by the model sum to one. However, computing this normalizing constant is often infeasible for high-dimensional data, making maximum likelihood estimation impractical.

To address this challenge, contrastive divergence is commonly employed as a practical approximation for training EBMs. Contrastive divergence seeks to minimize the difference between the model's distribution and the data distribution by iteratively updating the parameters of the neural network. By comparing samples generated from the model with observed data, contrastive divergence provides a tractable objective function for training EBMs without explicitly computing the normalizing constant.

Energy-based models (EBMs) offer a versatile framework with applications spanning across diverse domains, including computer vision, natural language processing (NLP), and more. In computer vision, EBMs play a crucial role in image generation and enhancement tasks. By training on large datasets of images, EBMs can effectively learn the underlying data distribution and generate new, realistic images. Moreover, EBMs excel in image denoising, inpainting, and super-resolution tasks by capturing intricate dependencies within the data. This capability is particularly valuable in medical imaging, where EBMs can generate synthetic medical images for training deep learning models. Synthetic images closely resembling real patient data augment the training data, enhancing the robustness and generalization of medical image analysis models.

In addition to computer vision, EBMs find applications in content generation for virtual reality (VR) and augmented reality (AR) environments. By training on a diverse range of visual data, including images, videos, and 3D models, EBMs can generate new content that seamlessly integrates with the virtual environment. This content generation capability enhances the immersive experience of VR and AR applications, creating realistic and engaging virtual worlds that captivate users.

Furthermore, EBMs hold promise in revolutionizing natural language processing (NLP) tasks, particularly in text generation and language modeling. Trained on large text corpora, EBMs can learn the intricate nuances and underlying distribution of natural language. This enables them to generate coherent and contextually relevant text, making them invaluable for applications such as dialogue generation, story generation, and machine translation. By generating human-like text, EBMs contribute to advancing the capabilities of NLP systems, enabling more natural and fluent interactions between humans and machines.

Overall, the versatility and flexibility of energy-based models make them indispensable tools for addressing challenging problems in data modeling and generation across various domains. By harnessing the power of EBMs, researchers and practitioners can unlock new possibilities in artificial intelligence and machine learning, driving advancements in technology and shaping the future of intelligent systems.

V. IMPLEMENTATION

In the practical implementation of energy-based models (EBMs) for image generation, several key steps and techniques are involved to ensure effective training and sampling.

Firstly, a neural network architecture is designed to serve as the energy function or score for the EBM. This neural network is typically composed of multiple layers, such as convolutional layers followed by fully connected layers, allowing it to capture complex patterns and dependencies within the data. In the context of image generation, the neural network is trained to assign low scores to realistic images and high scores to unrealistic ones, effectively learning the underlying data distribution.

To train the neural network, contrastive divergence is commonly utilized as the training objective. Contrastive divergence is a learning algorithm that approximates the gradient of the log-likelihood function without explicitly computing the normalizing constant. This enables efficient training of the neural network without the need for computationally expensive calculations, making it well-suited for large-scale datasets.

Once the neural network is trained, Langevin dynamics is employed for sampling new images from the trained EBM. Langevin dynamics is a stochastic process that iteratively adjusts the input image based on the gradient of the energy function. By simulating this process, Langevin dynamics can generate samples that closely resemble the training data distribution, allowing for the generation of realistic images.

Overall, the practical implementation of EBMs for image generation involves designing a neural network architecture, training the network using contrastive divergence, and sampling new images using Langevin dynamics. These techniques enable the generation of high-quality images that capture the underlying data distribution, demonstrating the effectiveness of EBMs in image generation tasks.

VI. CONCLUSION

In conclusion, this paper has delved into the intricacies of energy-based models (EBMs) and their practical implementation in the domain of image generation. By drawing inspiration from the Boltzmann distribution and leveraging the power of neural network architectures, EBMs offer a compelling framework for capturing the underlying data distributions in a variety of contexts.

Throughout this paper, we have highlighted the foundational principles behind EBMs, emphasizing their ability to assign low energy scores to likely observations and high energy scores to unlikely ones. This fundamental concept, rooted in the Boltzmann distribution, forms the basis of EBMs and allows them to model complex data distributions in a principled manner.

Despite the inherent challenges associated with EBMs, such as sampling new observations and addressing the intractable normalizing denominator, we have explored approximation techniques like contrastive divergence and Langevin dynamics. These techniques enable efficient training and sampling from EBMs, making them practical and scalable for real-world applications.

REFERENCES

[1] Zengyi Li, Yubei Chen, Friedrich T. Sommer. *2019*.Learning Energy-Based Models in High-Dimensional Spaces with Multi-scale Denoising Score Matching,

[2] Will Grathwohl, Kuan-Chieh Wang, Jörn-Henrik Jacobsen, David Duvenaud, Mohammad Norouzi, Kevin Swersky. *2020*. Your Classifier is Secretly an Energy Based Model and You Should Treat it Like One.

[3] Yang Song, Diederik P. Kingma. 2021. How to Train Your Energy-Based Models.

[4] Yee Whye Teh,Simon Osindero,Geoffrey E. Hinton. 2003. Energy-Based Models for Sparse Overcomplete Representations.

[5]Yann LeCun, Sumit Chopra, Marc'Aurelio Ranzato, Fu-Jie Huang. 2007. Energy-Based Models in Document Recognition and Computer Vision.

.