On the Hardness of Learning One Hidden Layer Neural Networks

Shuchen Li SHUCHEN.LI@YALE.EDU
Ilias Zadik ILIAS.ZADIK@YALE.EDU
Manolis Zampetakis EMMANOUIL.ZAMPETAKIS@YALE.EDU
Yale University

Editors: Gautam Kamath and Po-Ling Loh

Abstract

In this work, we consider the problem of learning one hidden layer ReLU neural networks with inputs from $\mathbb{R}^d$. It is well known due to (Klivans and Sherstov, 2009) that without further assumptions on the distribution $\mathcal{D}$, e.g., when $\mathcal{D}$ can be supported over the Boolean hypercube, learning even one-hidden layer neural networks is impossible (or "hard"[1]) for polynomial-time estimators under standard cryptographic assumptions. Given the success of neural networks in practice, a long line of recent work has attempted to study instead the canonical continuous input distribution case where $\mathcal{D}$ is the isotropic Gaussian, i.e., $\mathcal{D} = \mathcal{N}(0, I_d)$ which is also the setting that we follow in this work. Yet, despite a long line of research, it remains open whether there is a polynomial-time algorithm for learning one hidden layer neural networks when $\mathcal{D} = \mathcal{N}(0, I_d)$. It is known that a single neuron, i.e., zero hidden layer neural network, can be learned in polynomial time (Zarifis et al., 2024), while neural networks with more than two hidden layers are hard to learn (Chen et al., 2022). Nevertheless, the case of one hidden layer neural networks is not well understood.

In this paper we close this gap in the literature by answering the question of efficient learnability of neural networks with one hidden layer. We establish that under the CLWE assumption from cryptography (Bruna et al., 2021), learning the class of one hidden layer neural network with polynomial size under standard Gaussian inputs and polynomially small Gaussian noise is indeed computationally hard. Importantly, solving CLWE in polynomial time implies a polynomial-time quantum algorithm that solves the *worst-case* gap shortest vector problem (GapSVP) within polynomial factors, a widely believed hard task in cryptography and algorithmic theory of lattices (Micciancio and Regev, 2009). En route, we prove the hardness of learning Lipschitz periodic functions under standard Gaussian inputs and polynomially small Gaussian noise. This improves the previous result from (Song et al., 2021), which proved the hardness for polynomially small *adversarial* noise.

We also utilize the more general reductions between CLWE and classical LWE due to (Gupte et al., 2022). In particular, we show that if we assume the hardness of GapSVP with subexponential approximation factors $2^{O(d^\delta)}$ for $\delta \in (0,1)$, we can show the hardness of learning one hidden layer neural networks with polynomial size under Gaussian noise with 2^{-d^η} variance, where $\eta = \frac{\delta}{1+\delta} \in (0, 1/2)$. The current state-of-the-art algorithm for GapSVP is the celebrated Lenstra-Lenstra-Lovász (LLL) lattice basis reduction algorithm (Lenstra et al., 1982) which has approximation factor $2^{\Theta(d)}$. Hence, our results show that any polynomial time learning algorithm for one hidden layer neural networks for any variance of noise $\sigma^2 \geq 2^{-o(\sqrt{d})}$ would imply a major algorithmic breakthrough in the theory of lattices.

0. Extended abstract. Full version appears as (Li et al., 2024)
1. Following a standard convention, we refer to a computational task as "hard" if it is impossibile for polynomial-time methods.

References

Joan Bruna, Oded Regev, Min Jae Song, and Yi Tang. Continuous lwe. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, 2021.

Sitan Chen, Aravind Gollakota, Adam R Klivans, and Raghu Meka. Hardness of noise-free learning for two-hidden-layer neural networks. *arXiv preprint arXiv:2202.05258*, 2022.

Aparna Gupte, Neekon Vafa, and Vinod Vaikuntanathan. Continuous lwe is as hard as lwe & applications to learning gaussian mixtures. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1162–1173. IEEE, 2022.

Adam R Klivans and Alexander A Sherstov. Cryptographic hardness for learning intersections of halfspaces. *Journal of Computer and System Sciences*, 75(1):2–12, 2009.

Arjen Klaas Lenstra, Hendrik Willem Lenstra, and László Lovász. Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261(4):515–534, 1982.

Shuchen Li, Ilias Zadik, and Manolis Zampetakis. On the hardness of learning one hidden layer neural networks, 2024. URL https://arxiv.org/abs/2410.03477.

Daniele Micciancio and Oded Regev. *Lattice-based Cryptography*, pages 147–191. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. ISBN 978-3-540-88702-7. doi: 10.1007/978-3-540-88702-7_5. URL https://doi.org/10.1007/978-3-540-88702-7_5.

Min Jae Song, Ilias Zadik, and Joan Bruna. On the cryptographic hardness of learning single periodic neurons. In Marc'Aurelio Ranzato, Alina Beygelzimer, Yann N. Dauphin, Percy Liang, and Jennifer Wortman Vaughan, editors, *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, pages 29602–29615, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/f78688fb6a5507413ade54a230355acd-Abstract.html.

Nikos Zarifis, Puqian Wang, Ilias Diakonikolas, and Jelena Diakonikolas. Robustly learning single-index models via alignment sharpness. In *Forty-first International Conference on Machine Learning*, 2024.