Journal of Indian Acad. Math. ISSN: 0970-5120
Vol. , No. (Year) pp.

TITLE

Author[1], and Author[2]

Abstract. This is the place for abstract.

Keywords: xyz.

2010 AMS Subject Classification: xyx.

1. Introduction

Suppose X is a nonempty set...

2. Preliminaries and Notation

Consider a feedforward neural network $f_\theta : \mathbb{R}^{n_0} \to \mathbb{R}^{n_L}$ of depth L parameterized by $\theta = \{(W_\ell, b_\ell)\}_{\ell=1}^{L}$, where $W_\ell \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$ and $b_\ell \in \mathbb{R}^{n_\ell}$. The forward computation is defined recursively:

$$h_0 = x, \qquad z_\ell = W_\ell h_{\ell-1} + b_\ell, \qquad h_\ell = \sigma(z_\ell), \quad \ell = 1, \dots, L, \tag{1}$$

where $\sigma : \mathbb{R} \to \mathbb{R}$ is applied component-wise and $h_\ell \in \mathbb{R}^{n_\ell}$ denotes the activation vector at layer ℓ.

Definition 2.1 (Activation function properties)**.** *Let $\sigma : \mathbb{R} \to \mathbb{R}$ be...*

3. Main Results

Theorem 3.1 (Depth collapse)**.** *Let $f_\theta : \mathbb{R}^{n_0} \to \mathbb{R}^{n_L}$ be a ...*

Proof. We proceed by induction .... □

Corollary 3.2. *The hypothesis class ...*

Proposition 3.3 (Exponential gradient decay)**.** *For an L-layer network ....*

3.1. Subsection.

Proposition 3.4 (TanH gradient bound)**.** *Under the same hypotheses...*

4. Another Section

4.1. **Subsection.**

Theorem 4.1 (ReLU gradient preservation)**.** *In a ReLU network...*

Remark 4.2 (Computational cost)**.** *The exponential computation ...*

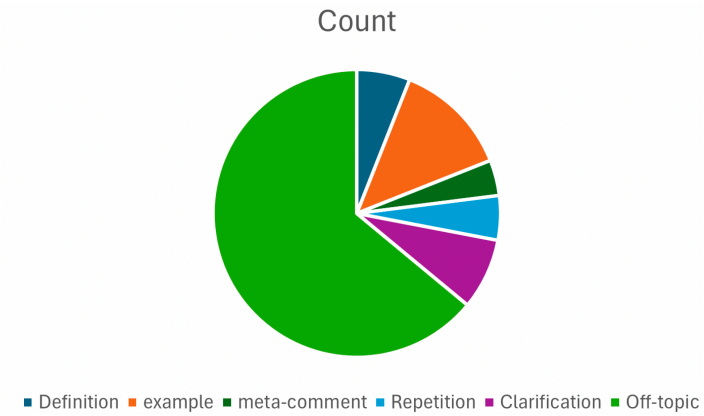


FIGURE 1. Name of Figure

5. Acknowledgement

The authors thank....

REFERENCES

[1] Z. Allen-Zhu, Y. Li, Z. Song, A convergence theory for deep learning via over-parameterization, in: *Proc. ICML*, PMLR **97**, 2019, pp. 242–252.
[2] Y. Bengio, P. Simard, P. Frasconi, Learning long-term dependencies with gradient descent is difficult, *IEEE Trans. Neural Netw.* **5**(2) (1994) 157–166.
[3] D.-A. Clevert, T. Unterthiner, S. Hochreiter, Fast and accurate deep network learning by exponential linear units (ELUs), *arXiv preprint* arXiv:1511.07289, 2015.
[4] X. Glorot, A. Bordes, Y. Bengio, Deep sparse rectifier neural networks, in: *Proc. AISTATS*, JMLR W&CP **15**, 2011, pp. 315–323.
[5] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, MIT Press, 2016.
[6] X. Glorot, Y. Bengio, Understanding the difficulty of training deep feedforward neural networks, in: *Proc. AISTATS*, JMLR W&CP **9**, 2010, pp. 249–256.
[7] K. He, X. Zhang, S. Ren, J. Sun, Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification, in: *Proc. IEEE ICCV*, 2015, pp. 1026–1034.
[8] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, *Neural Networks* **2**(5) (1989) 359–366.
[9] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature* **521**(7553) (2015) 436–444.
[10] V. Mnih et al., Human-level control through deep reinforcement learning, *Nature* **518**(7540) (2015) 529–533.
[11] P. Ramachandran, B. Zoph, Q. V. Le, Searching for activation functions, *arXiv preprint* arXiv:1710.05941, 2017.
[12] R. S. Sutton, A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed., MIT Press, 2018.

(Received, Month Date, Year)

(Revised, Month Date, Year)

[1]Department,
University,
Country.
Email[1] 1email@xyz.com
[1]Department,
University,
Country.
Email[2] 2email@xyz.com