

既存のニューラルネットワークにおける問題

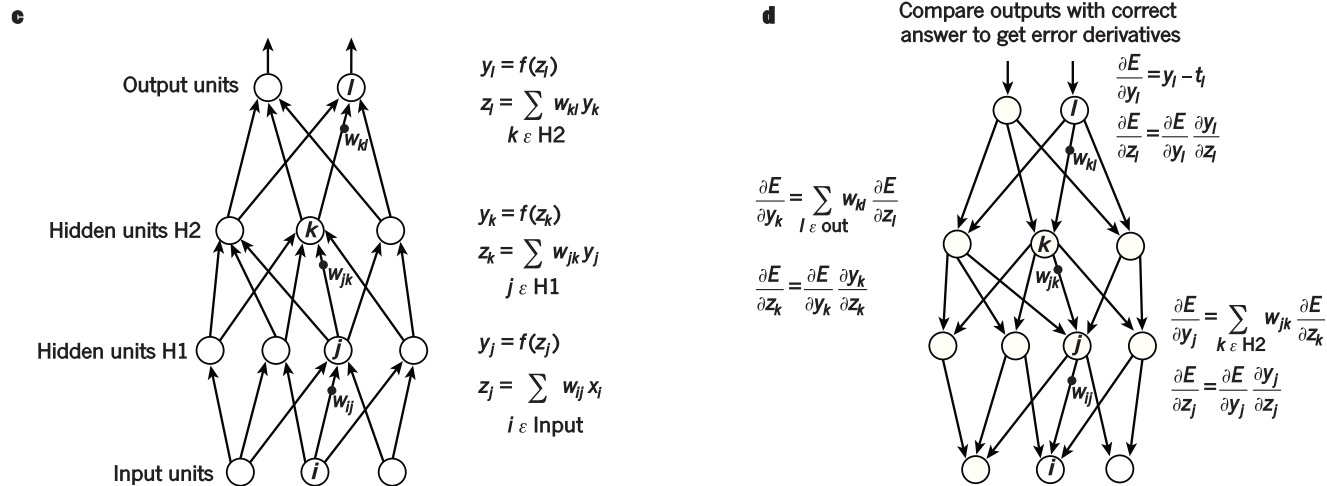
浅川伸一

既存のニューラルネットワークにおける問題

1. **勾配消失問題** the gradient vanishing problem (\implies **勾配爆発問題** the gradient exploding problem)
2. **信用割当問題** credit assignment problems (強化学習の意味での信用割当問題とは異なる)

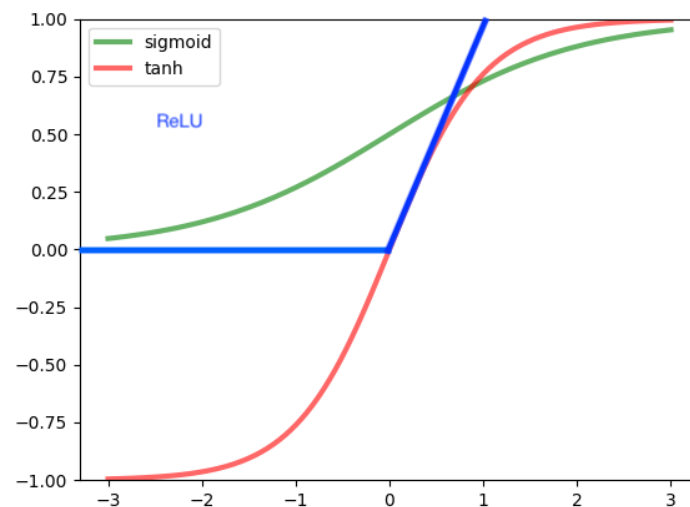
勾配消失問題 gradient vanishing problems

- チェインルール chain rule 合成関数の微分公式
- 活性化関数 activation functions
 - ロジスティックシグモイド関数 logistic sigmoid function (Rumelhart, Hinton, and Williams 1986)
 - ハイパータンジェント (LeCun et al. 1998)
 - 整流線形ユニット **ReLU** (Nair and Hinton 2010a, 2010b)



出典:(LeCun, Bengio, and Hinton 2015)

活性化関数

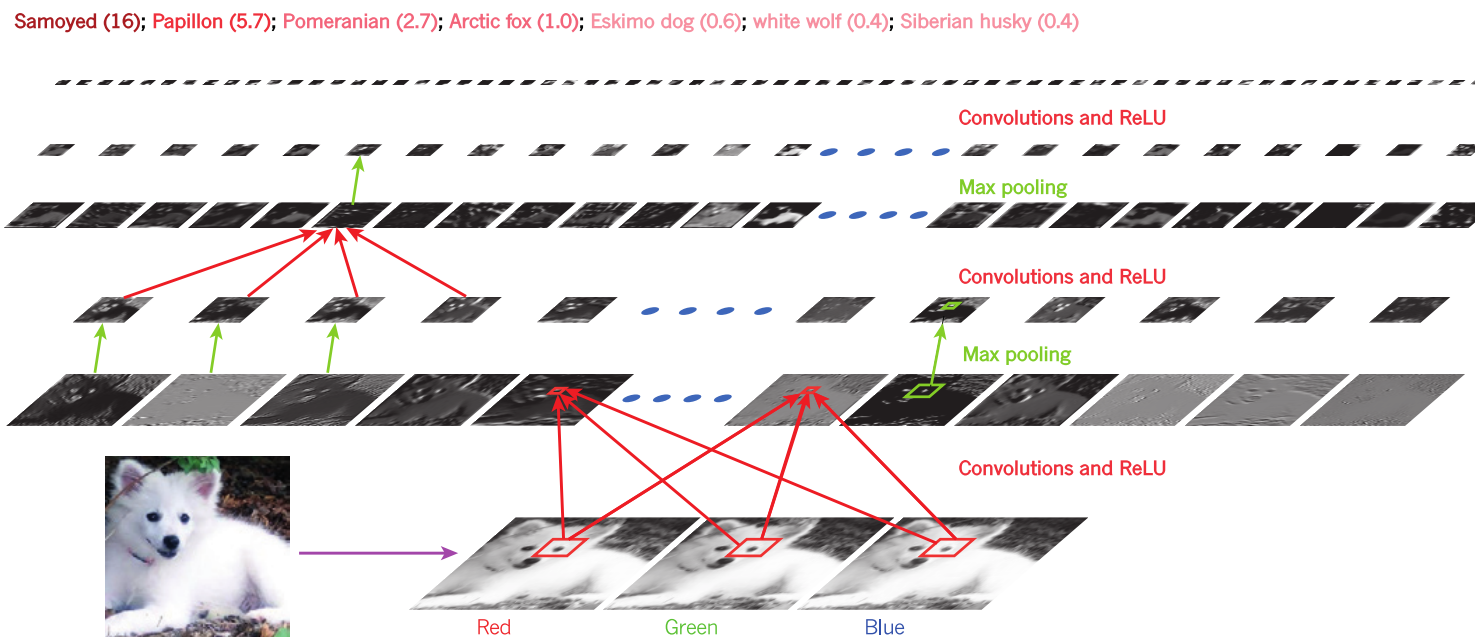


シグモイド関数, tanh, ReLU

- シグモイド関数 $y = \sigma(x) = (1 + e^{-x})^{-1}$, $f'(x) = y(1 - y)$
- ハイパータンジェント $y = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$, $f'(x) = 1 - y^2$
- ReLU: 整流線形ユニット **R**ectified **L**inear **U**nit $\text{ReLU}(x) = \max [0, x]$

信用割当問題 credit assignment problems

- 畳込み演算 convolution



(LeCun, Bengio, and Hinton 2015) より

まとめ

- 勾配消失問題: 活性化関数の工夫
- 信用割当問題: 畳み込みニューラルネットワーク

クイズ

活性化関数の中で、シグモイド関数と ReLU では、どちらが勾配消失問題が起こりにくいでしょうか？

クイズの答え

活性化関数の中で、シグモイド関数と ReLU では、どちらが勾配消失問題が起こりにくいでしょうか？

整流線形関数 ReLU

文献

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. “Deep Learning.” *Nature* 521: 436–44.

LeCun, Yann, Leon Bottou, Genevieve B. Orr, and Klaus-Robert Müller. 1998. “Efficient Backprops.” In *Neural Networks: Tricks and the Trade*, edited by Grégoire Montavon, Geneviève B. Orr, and Klaus-Robert Müller, 9–48. Berlin Heidelberg, Germany: Springer-Verlag. https://doi.org/10.1007/978-3-642-35289-8_3.

Nair, Vinod, and Geoffrey E. Hinton. 2010a. “Rectified Linear Units Improve Restricted Boltzmann Machines.” In *In Proceedings the 27th International Conference on Machine Learning (ICML)*, edited by Johannes Fürnkranz and Thorsten Joachims. Haifa, Israel: Omnipress. <http://www.icml2010.org/>.

———. 2010b. “Rectified Linear Units Improve Restricted Boltzmann Machines.” In *Proceedings of the 27th International Conference on Machine Learning (ICML)*, edited by Johannes Fürnkranz and Thorsten Joachims, 807–14. Haifa, Israel: Omnipress.

Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. 1986. “Learning Representations by Back-Propagating Errors.” *Nature* 323 (6088): 533–36. <https://doi.org/doi:10.1038/323533a0>.