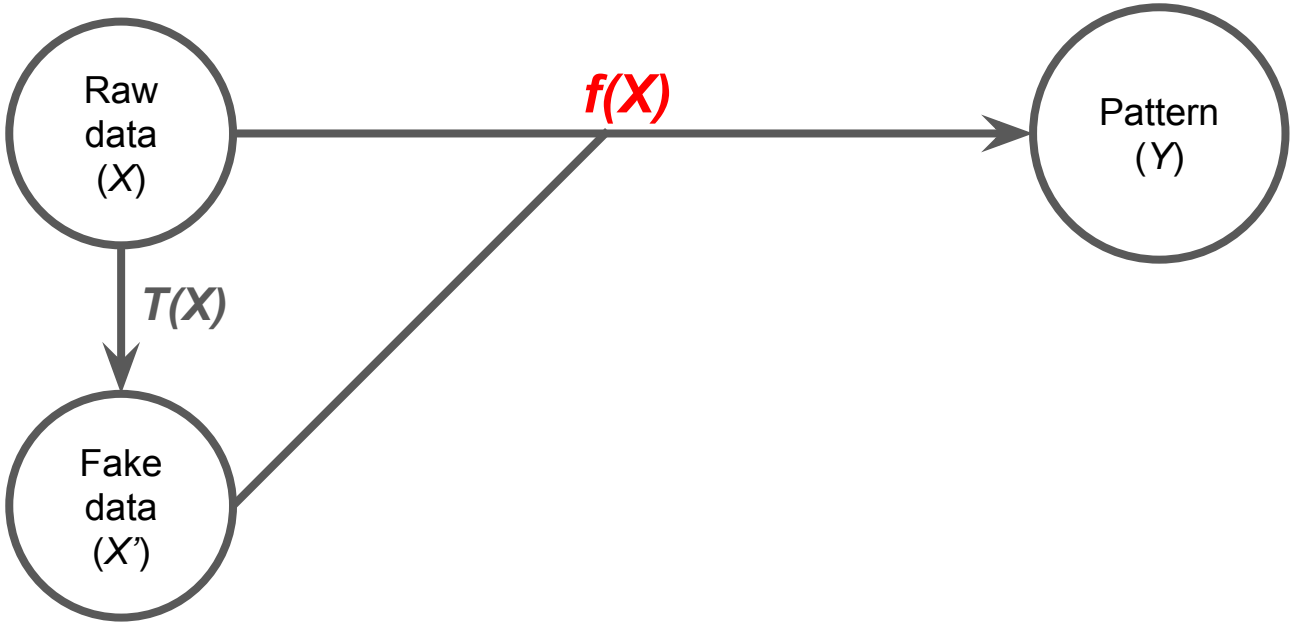
Advanced Machine Learning

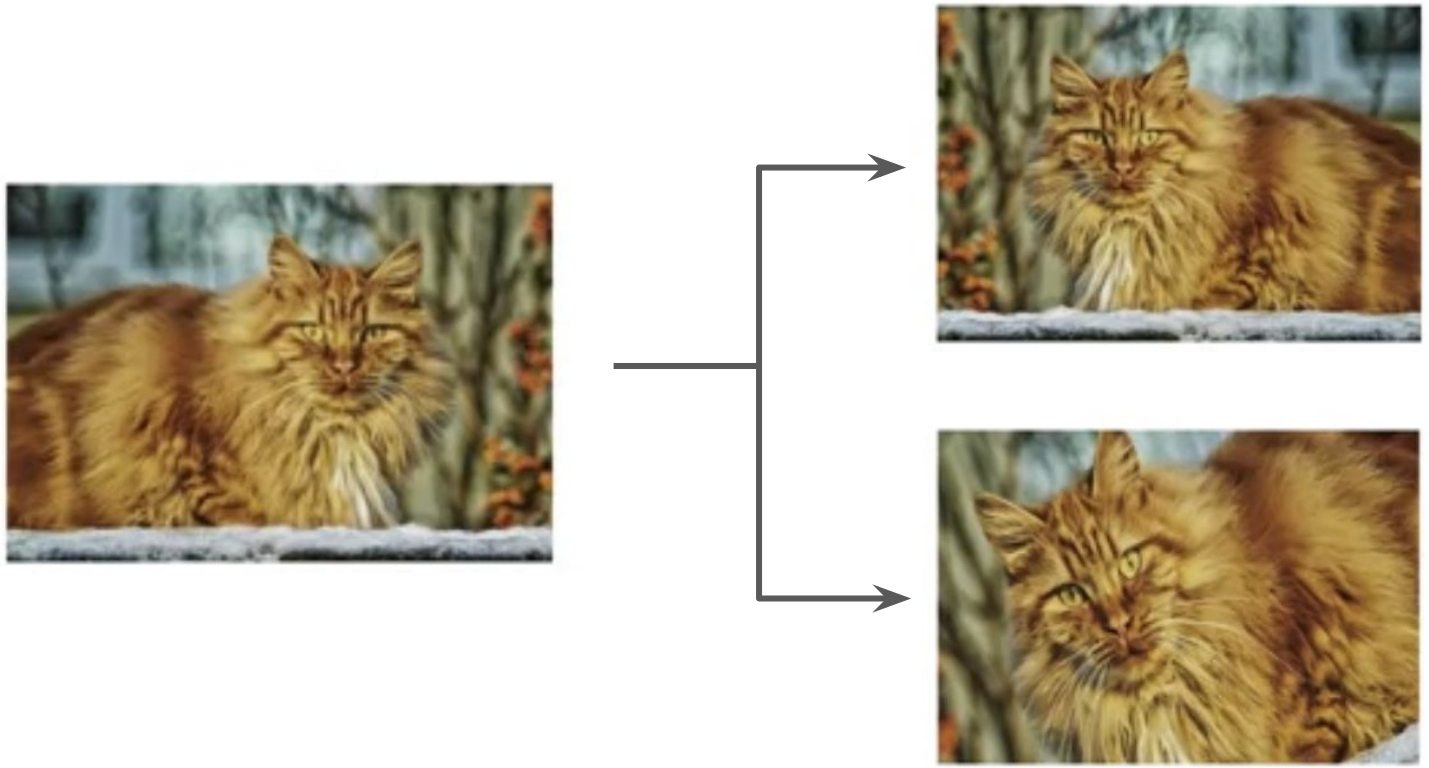
Likhit Nayak

Dataset Augmentation

Best way to make a machine learning model generalize better is to train it on more data - create fake data and add to the training set.

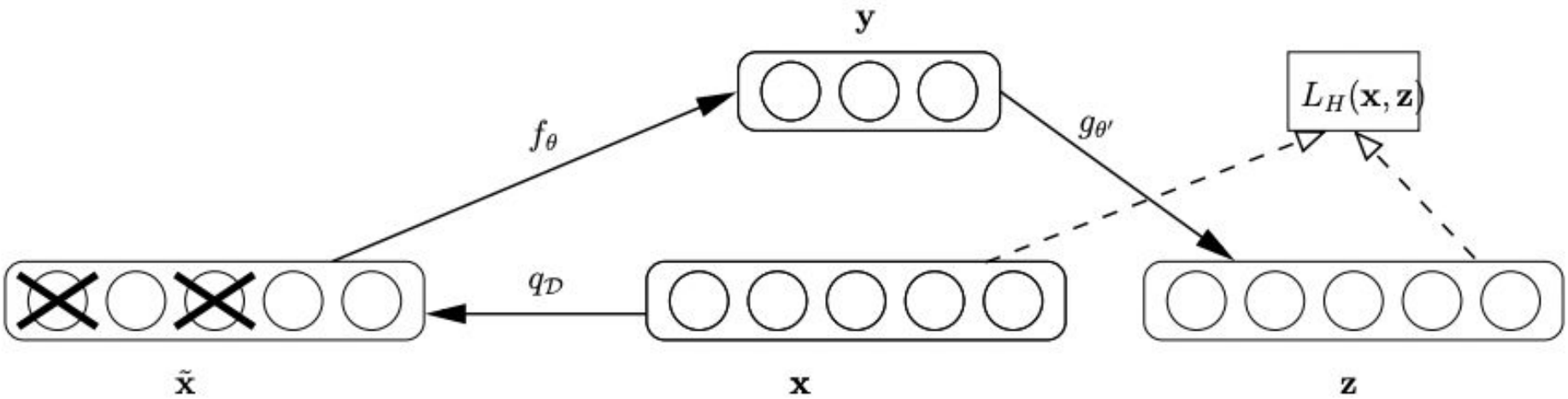


Dataset Augmentation



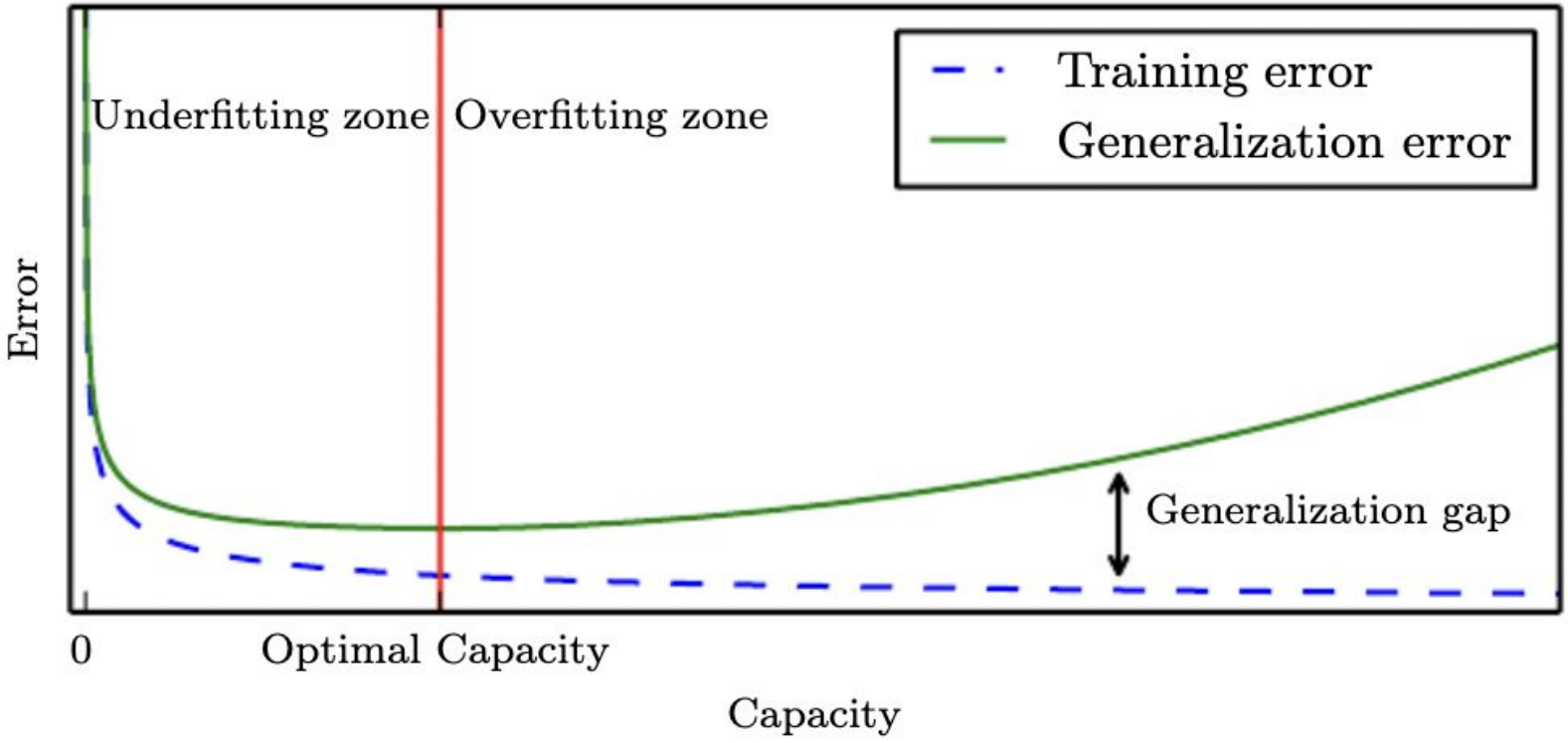
Dataset Augmentation

Dataset Augmentation - Injecting Noise

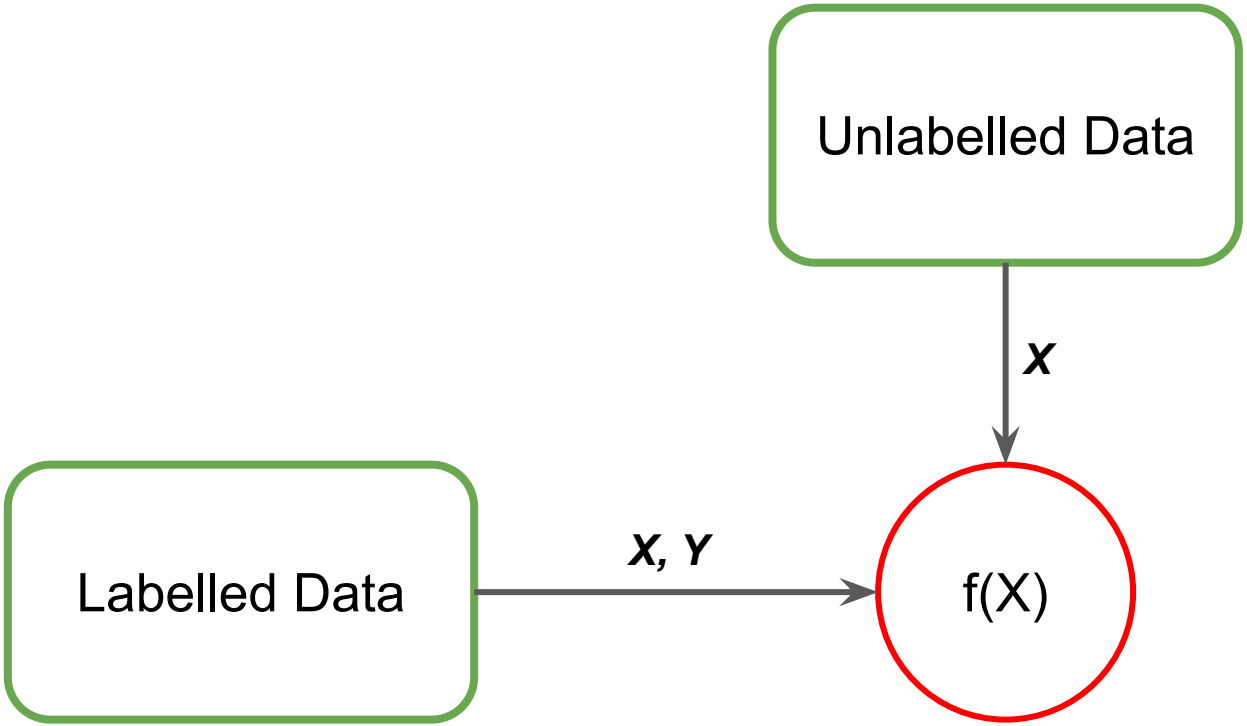


Vincent, Pascal, et al. "Extracting and composing robust features with denoising autoencoders." Proceedings of the 25th international conference on Machine learning. 2008.

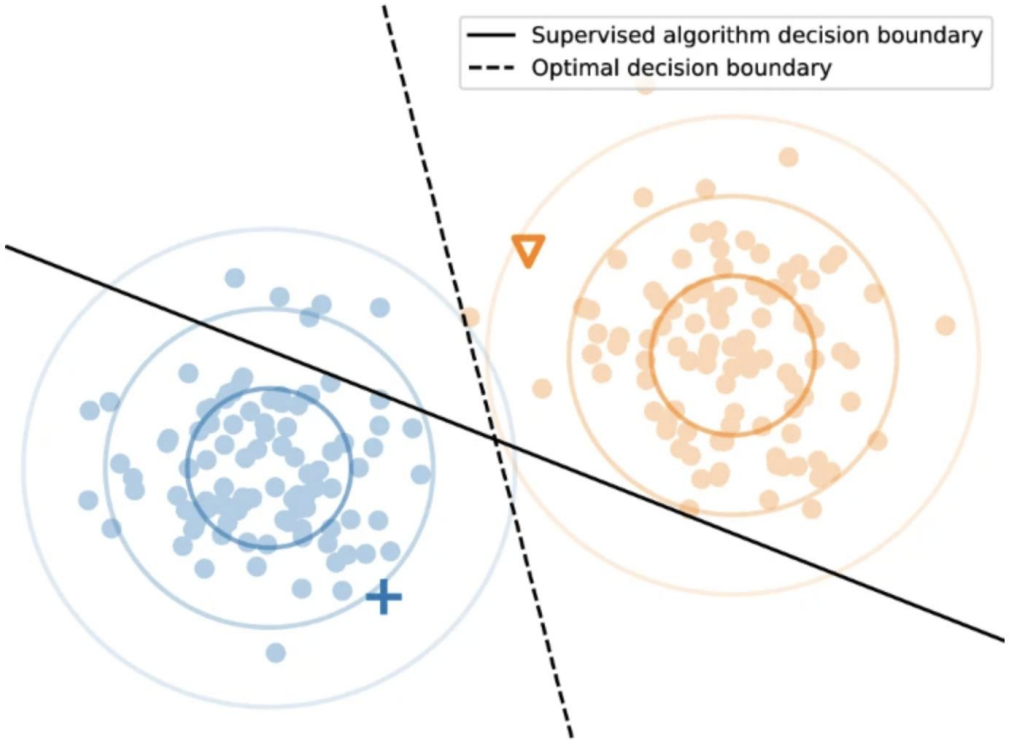
Early Stopping



Semi-supervised Learning



Semi-supervised Learning



Van Engelen, Jesper E., and Holger H. Hoos. "A survey on semi-supervised learning." Machine learning 109.2 (2020): 373-440.

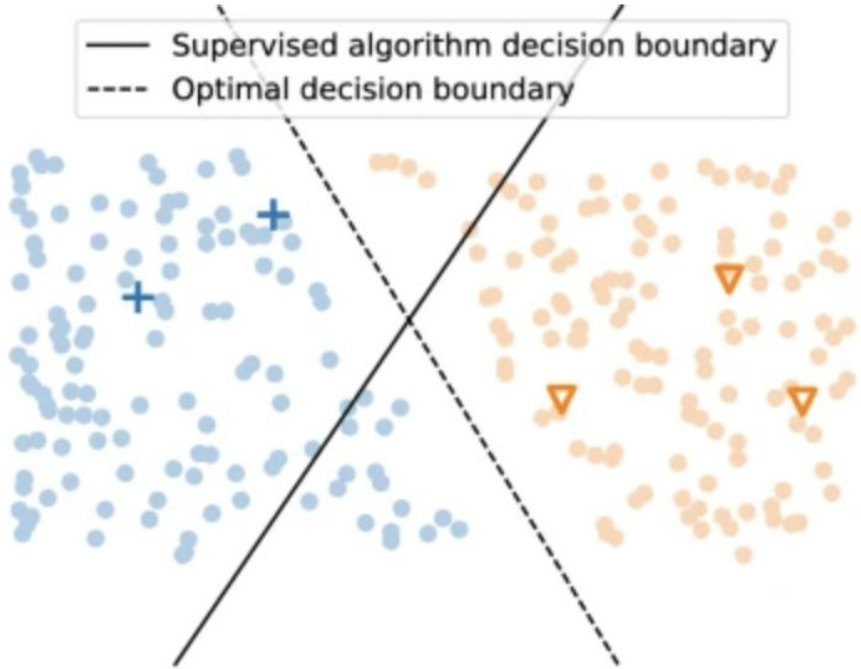
Semi-supervised Learning

Assumptions:

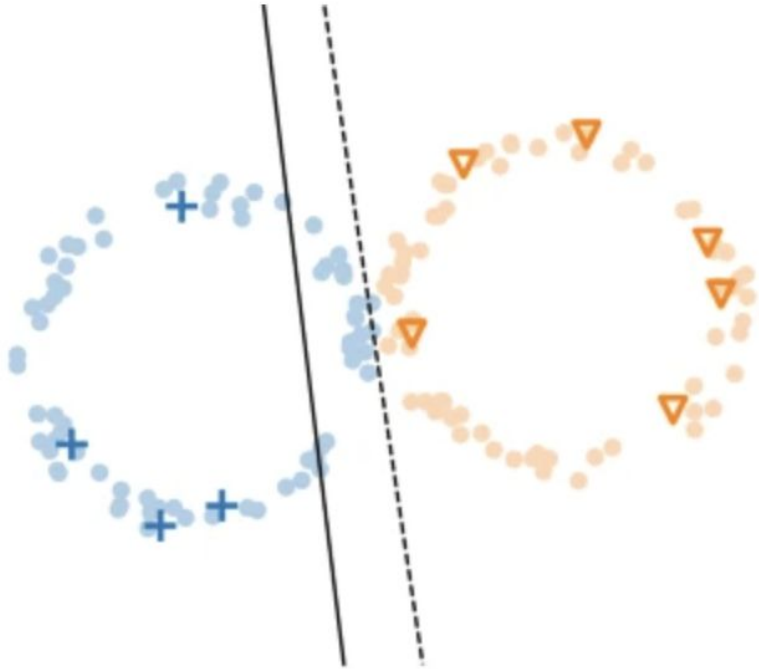
1. Smoothness assumption
   - If two samples $\boldsymbol{x}$ and $\boldsymbol{x'}$ are close in the input space, their labels $\boldsymbol{y}$ and $\boldsymbol{y'}$ should be the same
2. Low-density assumption
   - The decision boundary should not pass through high-density areas in the input space
3. Manifold assumption
   - The input space is composed of multiple lower-dimensional manifolds on which all data points lie
   - Data points lying on the same manifold have the same label

Van Engelen, Jesper E., and Holger H. Hoos. "A survey on semi-supervised learning." Machine learning 109.2 (2020): 373-440.

Semi-supervised Learning



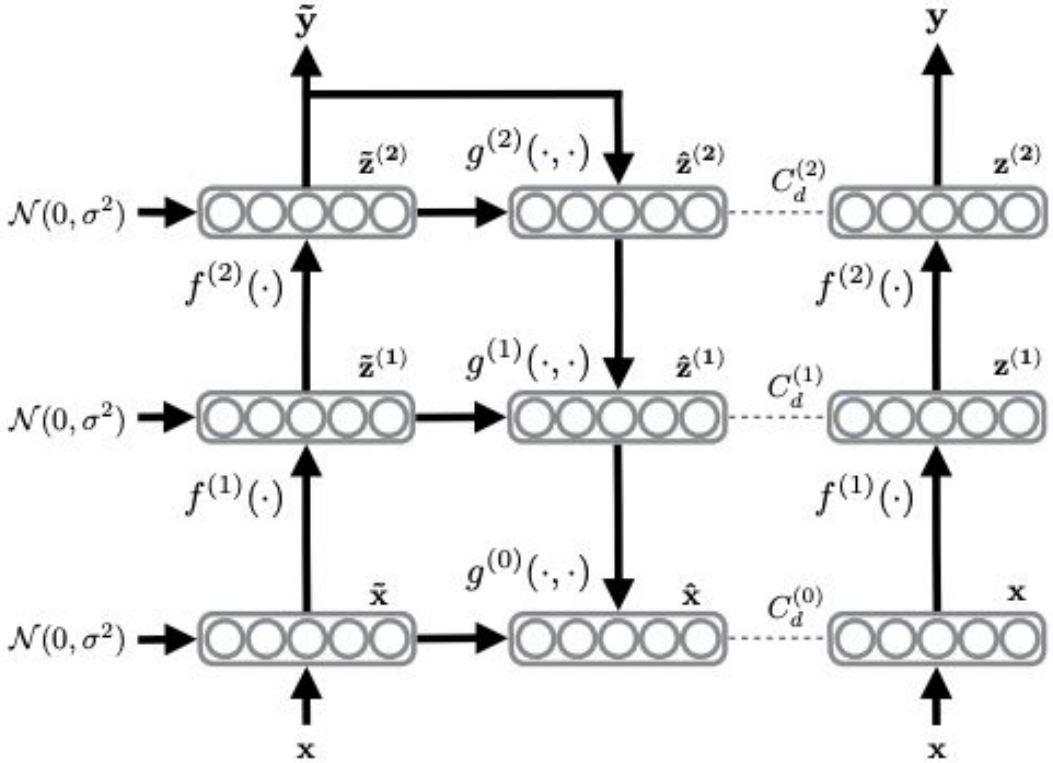
(a) Smoothness and low-density assumptions.



(b) Manifold assumption.

Van Engelen, Jesper E., and Holger H. Hoos. "A survey on semi-supervised learning." Machine learning 109.2 (2020): 373-440.

Ladder Networks



Rasmus, Antti, et al. "Semi-supervised learning with ladder networks." Advances in neural information processing systems 28 (2015).

Other Semi-supervised Networks

- Virtual adversarial training
  - Miyato, T., Maeda, S. I., Koyama, M., & Ishii, S. (2018). Virtual adversarial training: A regularization method for supervised and semi-supervised learning. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(8), 1979–1993.
- Π-model
  - Laine, S., & Aila, T. (2017). Temporal ensembling for semi-supervised learning. In International conference on learning, representations.

Multitask Learning

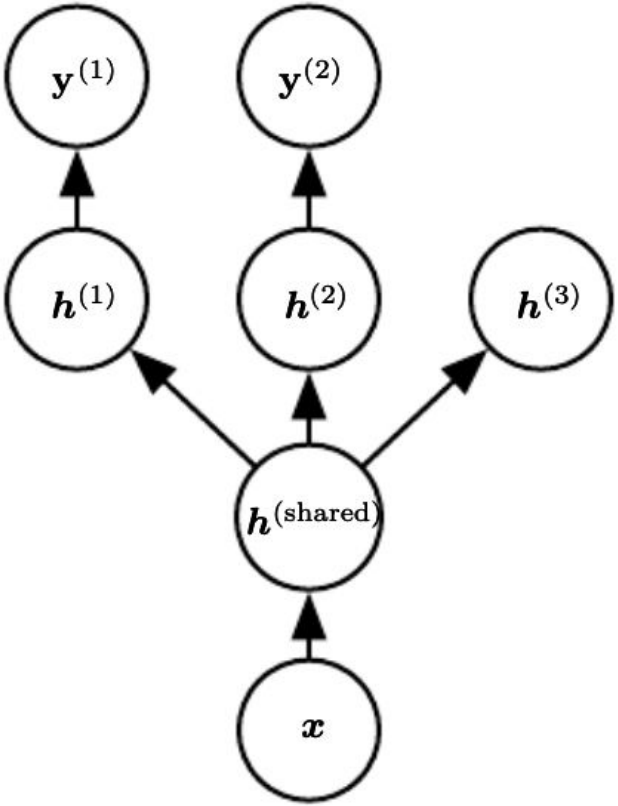
Assumption:

Among the factors that explain the variations observed in the data associated with different tasks, some are shared across two or more tasks.

Intuition:

In the same way that additional training examples put more pressure on the parameters of the model toward values that generalize well, when part of a model is shared across tasks, that part of the model is more constrained towards “good values”, thus yielding better generalization.

Multitask Learning



References

- Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep learning. MIT press, 2016.

Multitask Learning