Disentanglement (Reviews)

January 22, 2024

Contents

$$\text{Max } \mathbb{E}_{p(X)}\left[\mathbb{E}_{q(Z|X)}\left[\log p(X \mid Z)\right] - D_{KL}\left[q(Z \mid X) \parallel p(Z)\right]\right]$$

where Z is produced by encoder if it sees X. (X is label in output)

maximize the expectation of probability of producing X if decoder θ sees Z :

- p means if we fix model until X and produce Z many times, it's Gaussian
- Prior $p(Z) \sim \mathcal{N}(0, I)$ - please
- Z samples from distributions q_ϕ

$$r(X) = \mathbb{E}_q\left[q(Z \mid X)\right]$$

Assumptions on prior $p(Z)$ distribution.

1 Challenging Common Assumptions in the Unsupervised Learning of Disentangled Representations [1]

1.1 Literature

- **Goal** is to find a useful transformation/mapping/encoder from $\mathbf{X}$ to $\mathbf{z}$.
- **Key idea** of disentangled representation is the assumption that data can be described by only a few key factors that can reproduce data. (Representation dimension << data dimension)

- **Properties** of disentangled representation is to being compact, interpretable, independent, informative, & contain all the information of data.
- **Condition**: A disentangled representation is achieved if only a single factor changes in $\mathbf{X}$ when one of the elements in $\mathbf{z}$ changes (i.e., $\Delta\mathbf{X}_i \rightarrow \Delta\mathbf{z}_i$).
- **Loss Function**: Variational Auto Encoder (VAE) which is a probabilistic Autoencoder, where ϕ is encoder and θ is decoder, maximizes the expectation of producing $\mathbf{X}$ over all data points.

$$\max_{\phi,\theta} \quad \mathbb{E}_{p(\mathbf{X})}\left[\underbrace{\mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{X})}\left(\log p_\theta(\mathbf{X} \mid \mathbf{z})\right)}_{\text{Decoder}} - \underbrace{D_{KL}\left(q_\phi(\mathbf{z} \mid \mathbf{X}) \parallel p(\mathbf{z})\right)}_{\text{Encoder}}\right]$$

Where $\mathbf{z}$ is samples from Variational Distribution produced by encoder.

$$\underbrace{\max_{\phi,\theta} \quad \mathbb{E}_{p(\mathbf{X})}\left[\begin{array}{c}\underbrace{\mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{X})}\left(\log p_\theta(\mathbf{X} \mid \mathbf{z})\right)}_{\text{Decoder}} \\ \downarrow \\ \text{The probability of producing } \mathbf{X} \\ \text{if decoder sees } \mathbf{z} \text{ that produced under} \mathbf{X}\end{array}\right]}_{\text{Maximizes the similarity between } \mathbf{X} \text{ and } \hat{\mathbf{X}}} + \underbrace{\min_{\phi,\theta} \quad \mathbb{E}_{p(\mathbf{X})}\left[\begin{array}{c}\underbrace{D_{KL}\left(q_\phi(\mathbf{z} \mid \mathbf{X}) \parallel p(\mathbf{z})\right)}_{\text{Encoder}} \\ \downarrow \\ \text{The probabilistic distance between} \\ \text{prior and posterior distributions}\end{array}\right]}_{\substack{\text{Minimizes the dissimilarity between prior and} \\ \text{posterior by forcing the } q_\phi(\mathbf{z}|\mathbf{X}) \text{ to become } p(\mathbf{z})}}$$

- Disentanglement is a property of unsupervised models.

1.2 Contributions

- **Theorem 1** $\rightarrow$ Achieving disentangled representation is impossible without making assumptions (inductive bias) on model and data.
- Unlike methods prove aggregated posterior $q(\mathbf{z}_i)$ (that are sampled $(\mathbf{z}_i)$) are not correlated, experiments shows that representation dimensions $(\mathbf{z}_i)$ are correlated.
- Hyperparameters matter more than the model choice.
- No validation for the assumption that disentanglement is useful or it reduces the complexity of data.

1.3 Limits

- Assumption for prior distribution of latent variables.

1.4 Questions

- .

1.5 Future Research

- .

2 Terms

- **a**: .

3 Guide

- $\mathbf{N}$: Neuron
- a: Scalar
- $\mathbf{a}$: Vector
- $\mathbf{A}$: Matrix or Tensor
- $\mathcal{A}$: Set
- $\mathbb{R}$: Real Numbers Set
- $\mathbb{C}$: Complex Numbers Set
- C_L: Column of Matrix L
- R_L: Row of Matrix L

4 Cite

References

[1] Francesco Locatello, Stefan Bauer, Mario Lucic, Gunnar Raetsch, Sylvain Gelly, Bernhard Schölkopf, and Olivier Bachem. Challenging common assumptions in the unsupervised learning of disentangled representations. In international conference on machine learning, pages 4114–4124. PMLR, 2019.