Information Bottleneck Principle

- A theoretical framework for compression in neural networks.
- Balances:
  - **Compression**: Reduce information from x to z.
  - **Relevance**: Ensure z retains information about y.

Objective of Information Bottleneck

Minimize the following loss:

$$\mathcal{L} = I(x; z) - \beta I(z; y)$$

Where:

- $I(x; z)$: Mutual information between x and z.
- $I(z; y)$: Mutual information between z and y.
- β: Controls the trade-off.

Connection Between VAEs and Information Bottleneck

- VAEs implicitly optimize an information bottleneck objective.
- KL Divergence term in VAEs regularizes the latent space.