Optimization of the loss Function — VAE

$$\theta^*, \phi^* = \left\{ \underset{\theta \ \phi}{\arg\min} \right. \{ L(\theta, \phi) \}$$

$\Rightarrow$ variational Lower Bound or Evedience lower Bound <u>ELBO</u>.

From the loss function: KL divergence part is always positive.
therefore $L(\theta, \phi)$ is the lower bound for $\log P_\theta(x)$.

From Elbo Proof:

$$\log P_\theta(x) - \underbrace{\boxed{D_{KL}(Q_\phi(x|z) \| P(z|x))}}_{\geq 0} = -L(\phi, \theta).$$

<u>thus:</u> $L(\phi, \theta) \leq \log P_\theta(x)$ ELBO

By minimizing the loss function, we are technically increasing the likelihood of producing/generating the real data samples.

$$\theta^*, \phi^* = \underset{\theta \ \phi}{\arg\min} \ L(\theta, \phi)$$

Alternative optimization strategy:

$$\theta^* = \nabla_\theta \{ L(\theta, \phi) \} \qquad \phi = \text{constant}$$

$$\phi^* = \nabla_\phi \{ L(\theta, \phi) \}$$