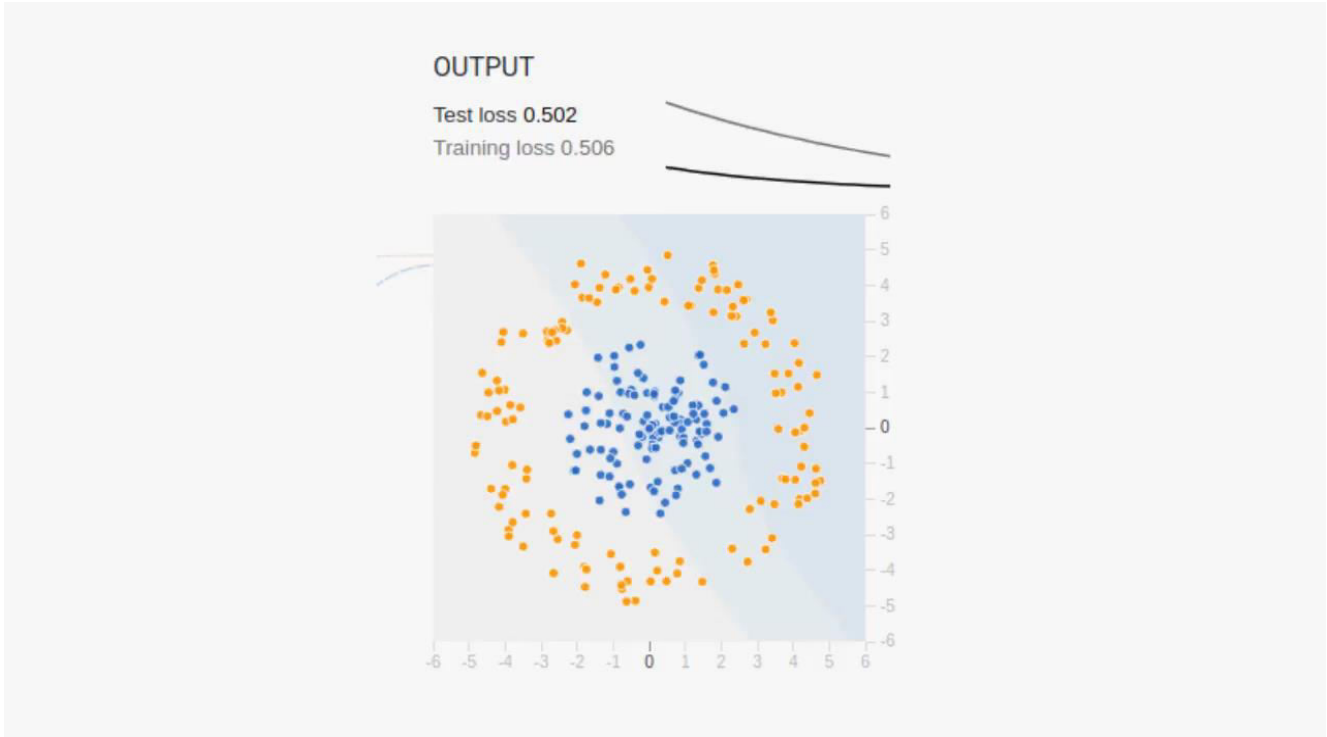
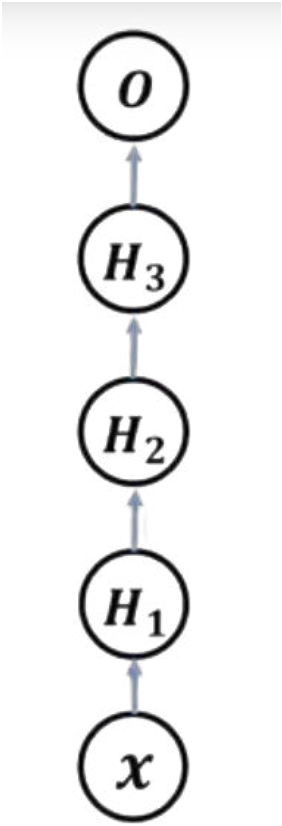
High Training Time

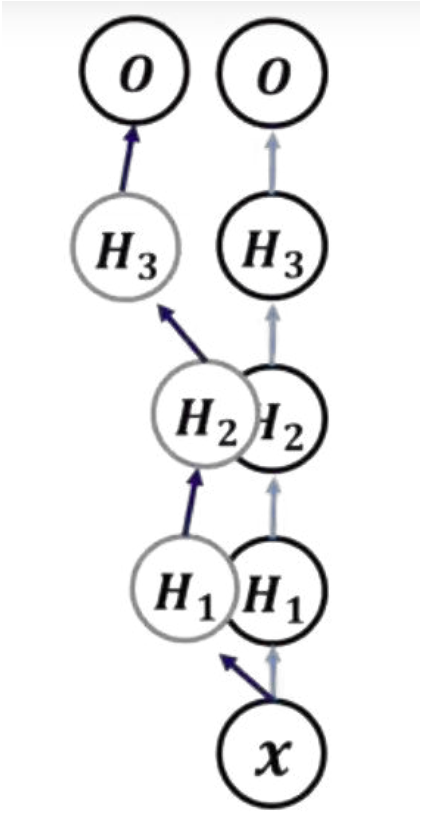


Neural Networks taking a lot of time to converge

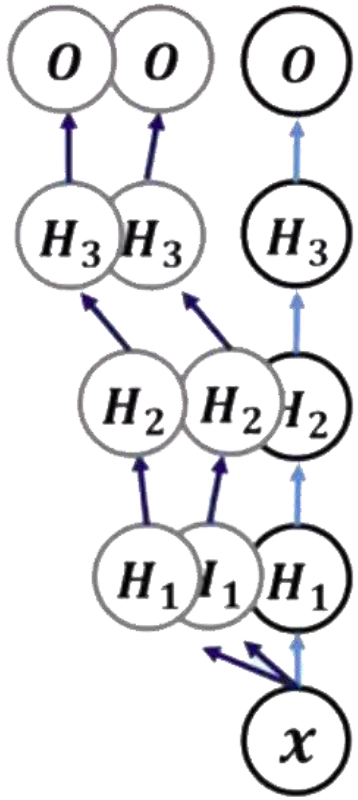
Reason: Change in Data Distribution across layers



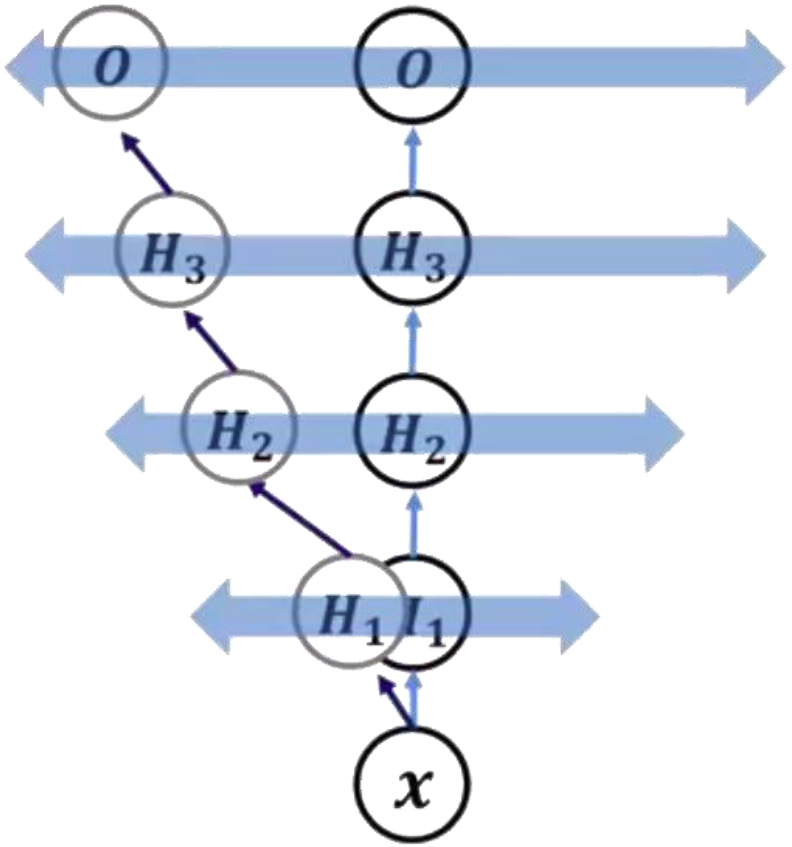
Problem : Change in Data Distribution across layers



Problem : Change in Data Distribution across layers

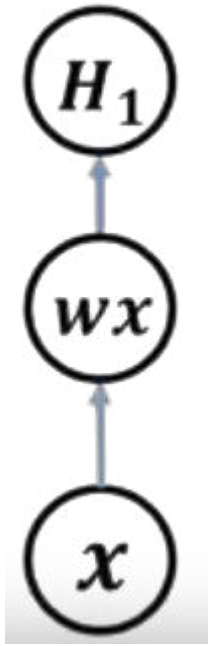


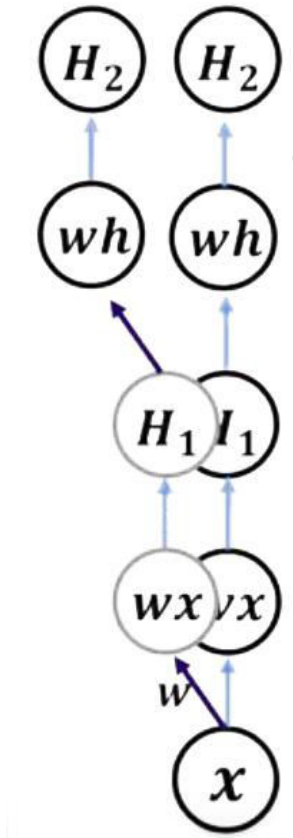
Problem : Change in Data Distribution across layers



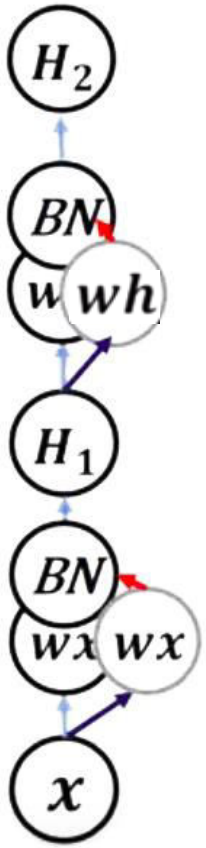
Internal Covariate Shift

Solution : BatchNormalization





Without BatchNorm



With BatchNorm

Analytics Vidhya
Learn everything about analytics

BatchNormalization : Technical Perspective

$$\mu_B = \frac{1}{m}\sum_{i=1}^{m} x^i$$

Batch Mean

$$\sigma_B^2 = \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu_B)^2$$

Batch Variance

$$\widehat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}$$

Normalize

$$y_i = \gamma\widehat{x} + \beta \equiv BN_{\gamma,\beta}\widehat{x}_i$$

Scale and Shift

Thank You