



**Data Science  
Academy**

[www.datascienceacademy.com.br](http://www.datascienceacademy.com.br)

**Deep Learning Frameworks**

**Lab**

**Batch Normalization com MXNet**



Em Machine Learning em geral, é muito comum normalizarmos a camada de entrada ajustando e dimensionando os dados.

Por exemplo, quando temos atributos (variáveis) com valores de 0 a 1 e alguns de 1 a 1000, devemos normalizá-los para acelerar o aprendizado (colocar todos os dados na mesma escala). Se a camada de entrada está se beneficiando da normalização, por que não fazer o mesmo com os valores nas camadas ocultas, que mudam o tempo todo? Ao fazer isso, poderíamos obter 10 vezes ou mais melhorias na velocidade de treinamento.

A normalização em lotes reduz a quantidade pela qual os valores das unidades ocultas mudam (mudança de covariância). Para explicar a mudança de covariância, considere como exemplo uma rede profunda para detecção de gatos. Treinamos nossos dados apenas nas imagens de gatos pretos. Portanto, se agora tentarmos aplicar essa rede a dados com gatos coloridos, não vamos nos sair bem. O conjunto de treinamento e o conjunto de previsão são imagens de gatos, mas diferem um pouco. Em outras palavras, se um algoritmo aprendeu algum mapeamento de X para Y e se a distribuição de X mudar, talvez seja necessário treinar novamente o modelo, tentando alinhar a distribuição de X com a distribuição de Y.

Além disso, a normalização em lote permite que cada camada de uma rede aprenda sozinha um pouco mais independentemente de outras camadas.

Ao usar Batch Normalization, podemos usar taxas de aprendizado mais altas porque a normalização em lote garante que não haja ativação muito alta ou muito baixa.

Batch Normalization também reduz a adaptação excessiva (overfitting), porque apresenta alguns efeitos de regularização. Semelhante ao Dropout, essa técnica adiciona algum ruído às ativações de cada camada oculta. Portanto, se usarmos a normalização em lote, usaremos menos dropouts, o que é uma coisa boa, pois não perderemos muitas informações. No entanto, não devemos depender apenas da normalização de lotes para regularização; devemos usá-lo em conjunto com o Dropout.

É o que faremos agora neste Lab. O Jupyter Notebook **“07-Batch-Normalization-MXNet”** ao final do capítulo tem todos os detalhes para você.

Boa aula.