



**Data Science
Academy**

www.datascienceacademy.com.br

Deep Learning II

Otimizando a Estrutura de uma
Rede Neural Artificial



Em Deep Learning I, aprendemos que nosso principal trabalho com redes neurais artificiais é ajustar os melhores pesos de uma rede neural, o que é feito através de algoritmos de treinamento, como algoritmos de Backpropagation por exemplo. Com efeito, esses algoritmos ajustam os pesos da rede neural de modo a reduzir o erro. Muitas vezes nos referimos aos pesos de uma rede neural como os parâmetros do modelo de rede neural. Alguns modelos de aprendizagem de máquina podem ter parâmetros diferentes dos pesos. Por exemplo, a regressão logística tem coeficientes como parâmetros.

Quando treinamos o modelo, os parâmetros de qualquer modelo de aprendizagem de máquina mudam. No entanto, esses modelos também possuem hiperparâmetros que não mudam durante o treinamento. Para as redes neurais, os hiperparâmetros especificam a arquitetura da rede. Exemplos de hiperparâmetros para redes neurais incluem o número de camadas ocultas e neurônios ocultos. Neste capítulo, examinaremos dois algoritmos que podem realmente modificar ou sugerir uma estrutura para a rede neural, otimizando assim o desempenho. O Pruning (ou “poda” em português) funciona analisando o quanto cada neurônio contribui para a saída da rede neural. Se a conexão de um neurônio com outro neurônio não afetar significativamente o resultado da rede neural, a conexão será “podada”. Através deste processo, as conexões e os neurônios que têm apenas um impacto marginal na saída, são removidos.

Outra técnica que apresentamos neste capítulo é a seleção do modelo (Model Selection). Enquanto o Pruning começa com uma rede neural já treinada, a seleção do modelo cria e treina muitas redes neurais com diferentes hiperparâmetros. O programa seleciona os hiperparâmetros que produzem a rede neural que atinge o melhor resultado de validação. Esta é uma técnica muito comum e poderosa, usada por exemplo em modelos baseados em método Ensemble, como o RandomForest.

As duas técnicas podem ser usadas para atingir o objetivo principal na construção de modelos de aprendizagem de máquina: a generalização. O modelo deve ser apresentar boa performance e precisão, sempre que for apresentado a um novo conjunto de dados.