

Machine Learning

Machine Learning

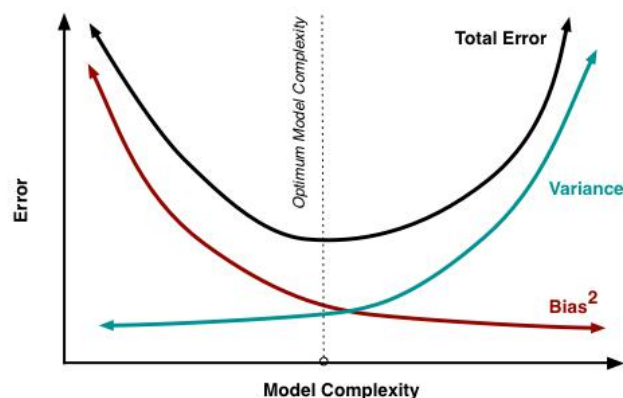
Regularização e Dropout

O objetivo do treinamento em redes neurais artificiais é obter uma rede que produza poucos erros no conjunto de treinamento, mas que também responda apropriadamente para novos padrões de entrada. A regularização é um método que busca melhorar a capacidade de generalização dos algoritmos de aprendizado, por meio de alguma restrição durante a fase de treinamento. A regularização ajuda a evitar o overfitting e melhora a generalização do modelo.

Em Deep Learning, é muito comum o modelo aprender demais sobre as peculiaridades nos dados, durante o treinamento e depois ter um baixo desempenho quando apresentado a novos dados. A regularização pode ser um remédio para este problema.

Regularização é qualquer modificação que nós fazemos no algoritmo de aprendizagem, com o objetivo de reduzir o erro de generalização, a fim de evitar o overfitting.

Existem muitas estratégias de regularização. Algumas colocam restrições extras em um modelo de aprendizado de máquina, como adicionar restrições nos valores dos parâmetros. Algumas acrescentam termos extras na função objetivo que podem ser considerados como correspondentes a uma restrição suave nos valores dos parâmetros. Se escolhidas cuidadosamente, essas restrições e penalidades extras podem levar a um melhor desempenho no conjunto de dados de teste. Às vezes, essas restrições e penalidades são projetadas para codificar tipos específicos de conhecimento prévio. Outras vezes, essas restrições e penalidades são projetadas para expressar uma preferência genérica por uma classe de modelo mais simples, a fim de promover a generalização.



No contexto de Deep Learning, a regularização de um estimador funciona através do aumento do bias para uma variância reduzida. Um regularizador efetivo é aquele que reduz significativamente a variância, sem aumentar excessivamente o bias (nós discutimos sobre bias e variância nos primeiros capítulos aqui do curso).

Na prática, um espaço de hipóteses excessivamente complexo não inclui necessariamente a função alvo ou o verdadeiro processo de geração de dados, ou mesmo uma aproximação de ambos. Nós quase nunca temos acesso ao verdadeiro processo de geração de dados, e nunca podemos saber com certeza se a família de modelos que está sendo estimada inclui o processo gerador ou não.

Os algoritmos de Deep Learning são tipicamente aplicados a domínios extremamente complicados, como imagens e sequências de áudio e texto, para os quais o verdadeiro processo de geração envolve essencialmente a simulação de todo o universo de dados. O que isto significa é que controlar a complexidade do modelo não é uma questão simples de encontrar o modelo do tamanho correto, com o número correto de parâmetros. Ao invés disso, podemos encontrar - e de fato em cenários práticos de aprendizagem profunda, quase sempre achamos - que o melhor modelo (no sentido de minimizar o erro de generalização) é um grande modelo que foi regularizado.

Portanto, seu objetivo em Deep Learning é encontrar um modelo que seja grande e profundo o suficiente para representar a complexidade nos dados e que possa ser aplicado a novos conjuntos de dados, com um bom desempenho. A regularização é uma das formas usadas para se alcançar esse objetivo.

O dropout é uma técnica de regularização criada por Geoffrey Hinton em 2012, portanto bem recente, que ajuda a mudar a saída dos neurônios de uma rede neural profunda, e que pode ser aplicado em qualquer camada das redes neurais profundas. O dropout desativa alguns dos neurônios da camada associada com alguma probabilidade p . Desativar um neurônio significa mudar o valor de saída para 0. No final, os neurônios que sofreram dropout tem os parâmetros reajustados, multiplicados por p (que é a probabilidade). O efeito de usar esse algoritmo é similar ao de fazer uma média de todos os possíveis modelos da rede neural que usam um subconjunto dos parâmetros disponíveis na camada afetada.