



**Data Science
Academy**

www.datascienceacademy.com.br

Deep Learning I

Métodos de Treinamento



A capacidade de aprender coisas novas é a principal característica de um cérebro humano. No caso de redes neurais artificiais, a aprendizagem é um processo de configuração de arquitetura de rede (a estrutura de conexões entre os neurônios) e dos pesos das ligações sinápticas (que afetam o coeficiente dos sinais) para obter uma solução eficaz para a tarefa em si. Normalmente a formação de uma rede neural é efetuada numa amostra de dados. O processo de formação segue um determinado algoritmo e com o passar do tempo, a reação de rede para os sinais de entrada devem melhorar.

Há três paradigmas principais de aprendizagem: supervisionada, não supervisionada e semi-supervisionada. No primeiro caso, os acertos para cada exemplo de entrada são conhecidos e os pesos tentam minimizar o erro. A aprendizagem não supervisionada permite categorizar as amostras através da explicação da estrutura interna e da natureza dos dados. Na formação semi-supervisionada ambas as abordagens acima são utilizadas.

Principais Regras da Aprendizagem da Rede Neural

Há quatro regras principais de aprendizagem baseadas na arquitetura de rede conectada com elas: correção de erros, a lei de Boltzmann, a regra de Hebb e o aprendizado competitivo. Vejamos uma breve explicação de cada uma delas.

1. Correção de Erros

Cada exemplo de entrada tem um valor de saída desejado especificado (valor objetivo), o qual pode incompatibilizar um valor real (previsão). A regra de aprendizagem da correção de erro se faz usando a diferença entre os valores objetivo e de previsão para ajuste direto dos pesos a fim de diminuir o erro. O treinamento é realizado somente no caso de um resultado errôneo. Esta regra de aprendizagem tem inúmeras variantes.

2. Regra de Boltzmann

A regra de Boltzmann é uma regra estocástica de aprendizagem por analogia com os princípios termodinâmicos. Isso resulta em coeficientes de peso dos neurônios ajustados de acordo com a distribuição probabilística desejada. Aprender a regra de Boltzmann pode ser considerado como um caso isolado de correção de um erro onde este signifique uma discrepância na correlação dos estados de dois modos.

3. Regra de Hebb

A regra de Hebb é o algoritmo de aprendizagem mais famoso das redes neurais. A ideia deste método é que se os neurônios de ambos os lados de uma sinapse ativarem simultaneamente e regularmente, então, a força dessa ligação sináptica aumenta. Uma particularidade importante aqui é que a mudança do peso sináptico depende apenas da atividade dos neurônios ligados a esta sinapse. Existem muitas variações a esta regra, que diferem em peculiaridades a modificação do peso da sinapse.

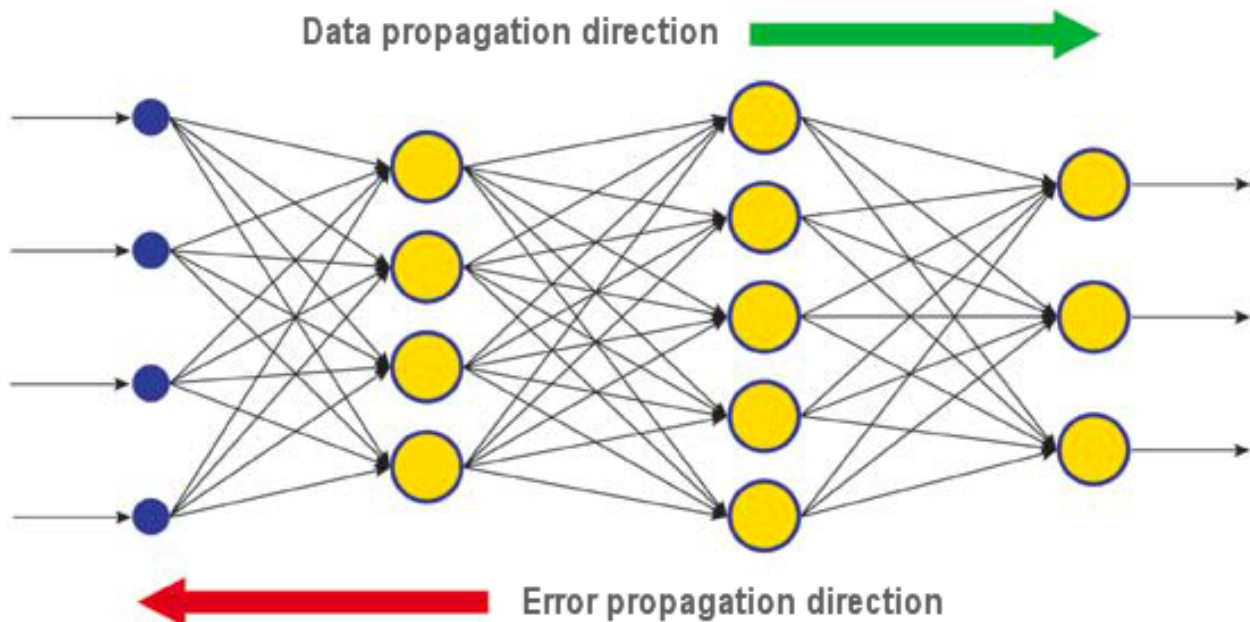
4. Aprendizagem Competitiva

Ao contrário da regra de aprendizagem de Hebb, onde um número de neurônios de saída pode ativar simultaneamente, aqui os neurônios de saída competem uns contra os outros. Um neurônio de saída com o valor máximo da soma dos pesos é o "vencedor" e "o vencedor leva tudo". As saídas de outros neurônios de saída são definidas para inativo. No aprendizado, apenas os pesos do "vencedor" serão modificados, visando o aumento da proximidade com a instância de entrada atual.

Backpropagation

Há uma série de algoritmos que abordam os diferentes problemas de aprendizagem. A retro-propagação (Backpropagation), um dos algoritmos modernos mais eficientes, é um deles. O princípio por trás disso é que a mudança de peso sináptico acontece em consideração ao gradiente local da função de erro.

A diferença entre as respostas corretas e reais de uma rede neural avaliada na camada de saída é propagada de volta, conforme imagem abaixo. Desta forma, cada neurônio pode definir a contribuição do seu peso para o erro cumulativo da rede. A regra de aprendizagem mais simples é o método de descida do gradiente, que é a mudança de peso sináptico proporcionalmente à sua contribuição para o erro cumulativo.





Certamente, este tipo de aprendizagem em rede neural não garante o melhor resultado de aprendizagem já que há sempre uma possibilidade de que o algoritmo entre em um mínimo local. Existem técnicas especiais que permitem obter a solução encontrada a partir de um ponto extremo local. Se depois de algumas aplicações desta técnica, a rede neural tem a mesma decisão, então pode-se concluir que a solução encontrada é mais provável de ser a ideal.

Dificuldades do Treinamento

A principal dificuldade do uso de redes neurais é a chamada "maldição da dimensionalidade". Quando as dimensões de entrada e o número de camadas são aumentados, a complexidade da rede e o tempo de aprendizagem estão crescendo exponencialmente e o resultado recebido nem sempre é o ideal.

Outra dificuldade do uso de redes neurais é que as tradicionais são incapazes de explicar como elas estão resolvendo as tarefas. Em alguns campos de aplicação, como na pesquisa de medicamentos, esta explicação é mais importante do que o próprio resultado. A representação do resultado interno é muitas vezes tão complexa que é impossível analisar, exceto os casos mais simples que normalmente não são de nenhum interesse.