



Data Science Academy

www.datascienceacademy.com.br

Deep Learning II

Grid Search Model Selection

A Grid Search é um algoritmo de força-bruta e tentativa e erro. Para esta técnica, você deve especificar cada combinação dos hiperparâmetros que você gostaria de usar. Você deve ser cauteloso em sua seleção porque o número de iterações de pesquisa pode crescer rapidamente. Esta especificação poderia ser a seguinte:

- Neurônios ocultos: 2 a 10
- Tamanho do passo (step size): 2
- Funções de ativação: tanh, sigmoid e ReLU

O primeiro item indica que o Grid Search deve tentar a contagem de neurônios ocultos entre 2 e 10 contando por 2, o que resulta no seguinte: 2, 4, 6, 8 e 10 (5 possibilidades totais). O terceiro item define que devemos também tentar as funções de ativação tanh, sigmoid e ReLU para cada contagem de neurônios. Este processo resulta em um total de quinze iterações porque cinco possibilidades vezes três possibilidades são quinze no total. Estas possibilidades estão listadas aqui:

Iteração # 1: [2] [sigmoid]
Iteração # 2: [4] [sigmoid]
Iteração # 3: [6] [sigmoid]
Iteração # 4: [8] [sigmoid]
Iteração # 5: [10] [sigmoid]
Iteração # 6: [2] [ReLU]
Iteração # 7: [4] [ReLU]
Iteração # 8: [6] [ReLU]
Iteração # 9: [8] [ReLU]
Iteração # 10: [10] [ReLU]
Iteração # 11: [2] [tanh]
Iteração # 12: [4] [tanh]
Iteração # 13: [6] [tanh]
Iteração # 14: [8] [tanh]
Iteração # 15: [10] [tanh]

Cada conjunto de possibilidades é chamado de eixo. Estes eixos rodam através de todas as combinações possíveis antes de terminarem. A maioria dos frameworks de Deep Learning permite dois tipos de eixos. O primeiro tipo é um intervalo numérico com uma etapa. O segundo tipo é uma lista de valores, como as funções de ativação acima.

Como você pode perceber, o Grid Search pode resultar rapidamente em um grande número de iterações. Considere se você deseja procurar o número ótimo de neurônios ocultos em cinco camadas, onde você permitiu até 200 neurônios em cada nível. Esse valor seria igual a 200 multiplicado por si mesmo cinco vezes, ou 200 a quinta potência. Este processo resulta em 320 bilhões de iterações. Como cada iteração envolve o treinamento de uma rede neural, as iterações podem levar minutos, horas ou até dias para serem executadas.

Ao realizar Grid Search, o processamento de multi-threading pode ser benéfico. A execução das iterações através de um pool de threads pode acelerar a busca. O pool de threads deve ter um tamanho igual ao número de núcleos no dispositivo (CPU ou GPU). Essa característica permite que uma máquina com oito núcleos trabalhe em oito redes neurais ao mesmo tempo. O treinamento dos modelos individuais deve ser de thread única quando você executa as iterações simultaneamente. Muitas estruturas usarão todos os núcleos disponíveis para treinar uma única rede neural. Quando você tem um grande número de redes neurais para treinar, você sempre deve treinar várias redes neurais em paralelo, executando-as uma vez para que cada rede use os núcleos da máquina.

No curso de Programação Paralela em GPU, o terceiro da Formação IA, estudamos como executar threads em paralelo na GPU.