

Redes Neurais Convolucionais 2

A tarefa em questão abordou o acompanhamento de algumas seções de um curso chamado “Deep Learning: Convolutional Neural Networks in Python”. Tal curso trouxe de forma aprofundada diversos conteúdos que já haviam sido estudados superficialmente no bootcamp. O estudo de redes neurais convolucionais (CNN) tem se mostrado de extrema importância para o aprendizado de máquina, abrindo a possibilidade de criar modelos que interpretam textos e imagens. Alguns dos principais conceitos estudados no curso estão descritos abaixo:

Redes neurais convolucionais

Uma Rede Neural Convolucional é uma arquitetura de aprendizado profundo projetada para processar e analisar dados que possuem uma estrutura de grade, como imagens. Diferentemente de redes neurais tradicionais, as CNNs são especialmente eficientes na extração de padrões espaciais e de hierarquias complexas. A principal ideia é a aplicação de operações de convolução para detectar características locais em diferentes partes da entrada. A parte matemática envolve a aplicação de filtros (kernels) sobre a entrada, seguida pela aplicação de funções de ativação.

Processamento de linguagem natural

No contexto do processamento de linguagem natural (PLN), as CNNs são frequentemente utilizadas para analisar dados textuais. Os embeddings são uma técnica crucial que converte palavras ou frases em vetores de números. Esses vetores capturam relações semânticas e são usados como entrada para as CNNs. A aplicação de convoluções em embeddings permite a extração de características relevantes para a compreensão do contexto textual, identificando padrões e relações.

CNNs para textos são especialmente úteis na identificação de sequências de palavras significativas, tornando-as eficazes para tarefas como classificação de sentimentos, análise de sentimento e até mesmo tradução automática.

Funções de Perda

As funções de perda são fundamentais na avaliação do desempenho de um modelo durante o treinamento. O Mean Squared Error (Erro Quadrático Médio) é comumente usado para problemas de regressão, enquanto o Binary Cross Entropy é aplicado a problemas de classificação binária. Já o Categorical Cross Entropy é preferido em problemas de classificação multiclasse, onde cada exemplo pertence a uma única classe.

Método do Gradiente

O método do gradiente é uma técnica crucial na otimização de modelos de aprendizado de máquina. Em resumo, envolve a atualização iterativa dos pesos do modelo na direção oposta ao gradiente da função de perda em relação aos pesos. Esse processo visa minimizar a função de perda, ajustando os pesos para otimizar o desempenho do modelo. O tamanho do passo, conhecido como taxa de aprendizado, é um parâmetro crítico nesse processo, influenciando a convergência do modelo durante o treinamento.

Em conclusão, o curso "Deep Learning: Convolutional Neural Networks in Python" proporcionou uma compreensão abrangente e aprofundada sobre redes neurais convolucionais e processamento de linguagem natural. A aplicação prática desses conhecimentos em projetos desenvolvidos durante o curso está disponível no meu perfil do GitHub.