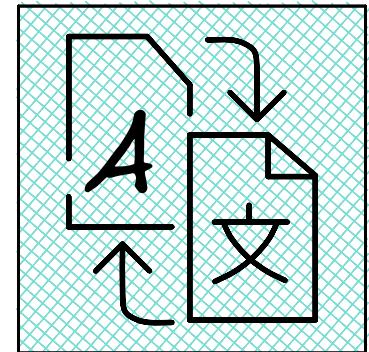
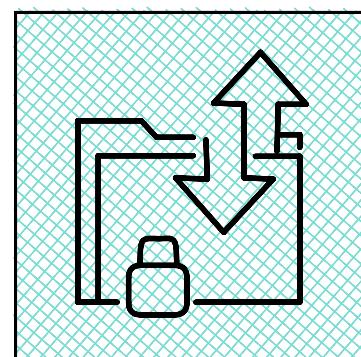


Transfer Learning

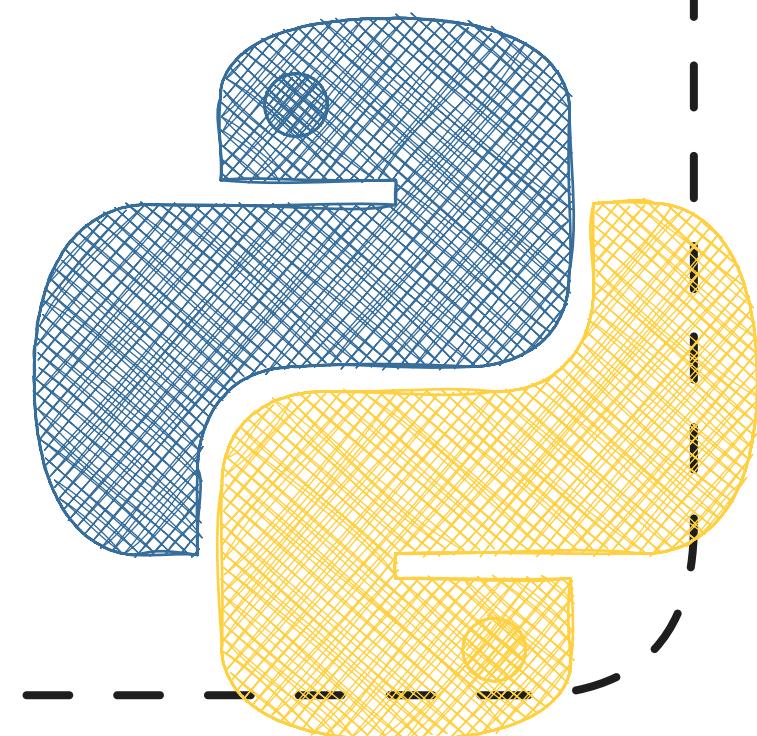
visão intuitiva



Técnica que aproveita o conhecimento previamente adquirido

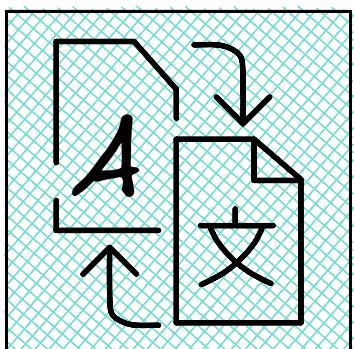


Transfer



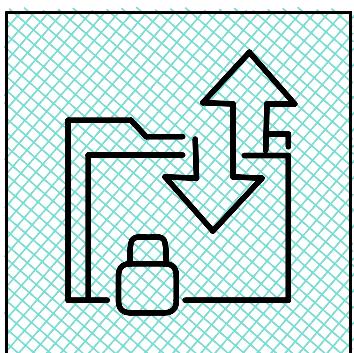
Reutilização direta

Feature Extraction



Um modelo pré-treinado é usado como extrator fixo de embeddings

Principal característica

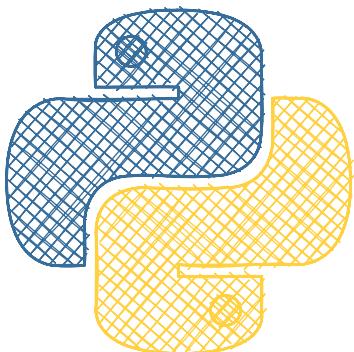


Camadas de convolução permanecem fixas

Apenas o classificador final é treinado

Exemplo:

usar a ResNet50 pré-treinada no ImageNet e treinar apenas uma camada densa que classifica diferentes tipos de folhas



Novo contexto

Fine-Tuning

Parte ou todo o modelo pré-treinado é reajustado

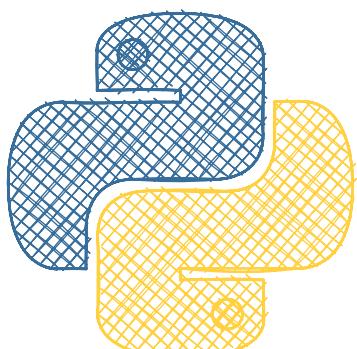
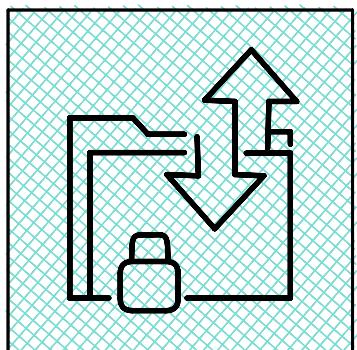
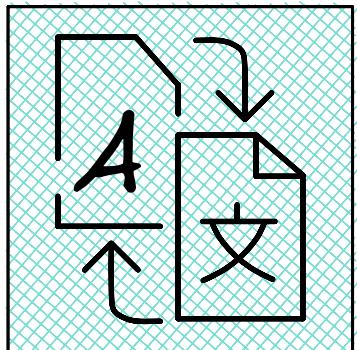
Principal característica

Permite que o modelo se adapte ao novo domínio

É mantido o conhecimento geral aprendido

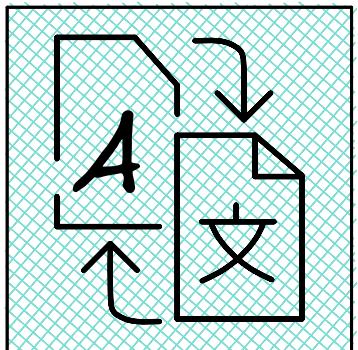
Exemplo:

ajustar o BERT para análise de sentimentos em texto em português



Compressão de conhecimento

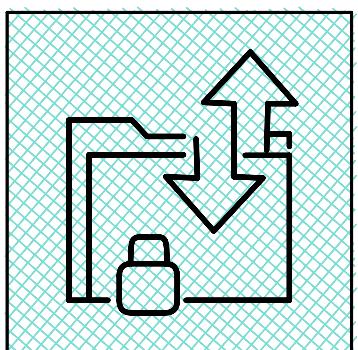
Knowledge Distillation



Transfere o conhecimento de um modelo grande para um modelo pequeno

Principal característica

O aluno aprende a imitar as previsões do professor



Mantendo desempenho semelhante com menos parâmetros

Exemplo:

distilar o GPT-4 em um modelo para rodar em dispositivos móveis

