

EEE950 - Redes Neurais Artificiais: Teoria e Aplicações

Exercício CNN - Implementação de uma CNN

José Geraldo Fernandes
Escola de Engenharia
Universidade Federal de Minas Gerais
Belo Horizonte, Brasil

I. INTRODUÇÃO

Treinou-se o modelo com os mesmos parâmetros da arquitetura implementada e verificou-se seu desempenho pobre.

A Figura 1 mostra a evolução da acurácia avaliada no conjunto de treinamento e validação.

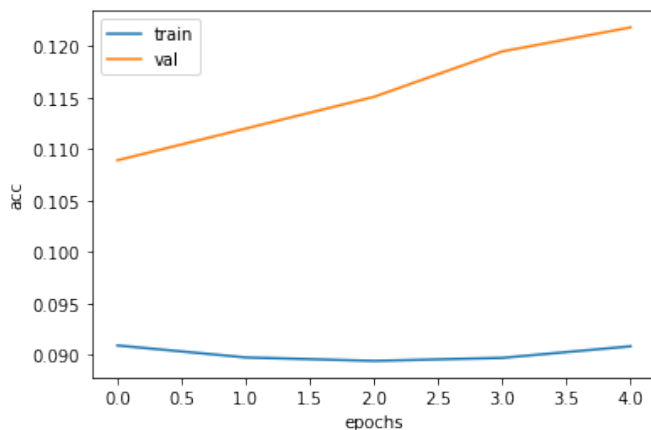


Figura 1. Curvas de acurácia do modelo original durante o treinamento.

A acurácia avaliada no conjunto de teste foi de 0.1147.

II. DISCUSSÃO

Para buscar melhorar a arquitetura do modelo deve-se, primeiro, analisar seu desempenho. Note que, tampouco no conjunto de treinamento houve aprendizado e a acurácia manteve-se baixa por todas as épocas.

Esse experimento sugere que a capacidade do modelo não é suficiente para a tarefa. Isso pode ser facilmente regulado a partir de parâmetros modularizados da arquitetura.

III. MÉTODO

No intuito de aumentar a capacidade do modelo, aumentou-se significativamente a quantidade de filtros, 16 e 32, removeu-se a primeira camada de *Dropout*, não é interesse regularizar nesse momento, e dobrou-se o número de épocas para treinamento.

Para compensar o aumento do tamanho do modelo em tempo de treinamento, dobrou-se o *batch size* e também

alterou-se o otimizador, para Adam [1] que é mais convencional.

IV. RESULTADOS

O mesmo experimento foi realizado com a arquitetura adaptada. A Figura 2 mostra a evolução, novamente, da acurácia avaliada no conjunto de treinamento e validação.

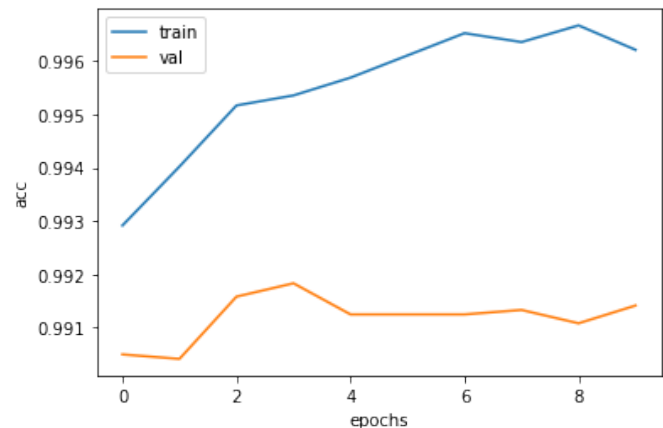


Figura 2. Curvas de acurácia do modelo adaptado durante o treinamento.

A acurácia avaliada no conjunto de teste foi de 0.9931.

V. CONCLUSÃO

Verificou-se a hipótese que aumentando a capacidade do modelo seria suficiente para melhorar seu desempenho.

Note que não só a acurácia está alta como também responde, aumenta gradativamente, com o treinamento.

O resultado no conjunto de teste também foi suficiente.

REFERÊNCIAS

- [1] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.