

Relatório Projeto 3

Aluno: João Guilherme Madeira Araújo

Número USP: 9725165

Relatório do Projeto 3:

O exercício pediu a implementação de um modelo Rede PCA (Principal Component Analysis) para compressão de coordenadas do dataset wine, em seguida comparamos o aprendizado de uma rede MLP (multilayer perceptron) no dataset original e no dataset comprimido. O framework utilizado na implementação do MLP foi o Tensorflow com Keras, o programa é chamado `adaptative_pca.py`.

Datasets

Foi usado o dataset wine, que contém 177 vinhos em 3 classes distintas, cada vinho tem 13 características como teor alcoólico, percentual de ácido málico e outros. 70% do dataset foi usado para treino e 30% foi usado para validação.

A Rede PCA adaptativa

A ideia da rede é usar uma variação da regra de Hebb chamada de regra de Oja para computar os Principal Components do nosso conjunto de dados. Usando apenas a regra de Oja conseguimos encontrar apenas o maior principal component, por isso usamos uma arquitetura um pouco diferente chamada rede de Rubner-Tava, que além dos pesos da regra de Oja tem pesos "laterais" treinados com uma regra anti-hebbiana usados para decorrelacionar a saída da rede, permitindo computar todos os componentes principais dos nossos dados. A rede MLP tem duas camadas com função de ativação ReLU para a camada escondida e softmax para a camada de saída.

Treinamento

O treinamento durou 60 épocas com função de perda entropia cruzada e método de otimização Adam. Para treinar a rede PCA foi usado uma learning rate decrescente e o treinamento foi feito até a convergência da rede.

Conclusões

O MLP com a PCA demorou menos para treinar, uma vez que ele só tinha 5 features de entrada. Além disso ele alcançou uma melhor acurácia durante o teste e demonstrou uma generalização bem melhor que no caso da MLP sem PCA que está claramente sofrendo de overfitting dos dados. Algo bastante interessante de se notar é que a MLP com PCA alcançou 10% mais acurácia que a MLP pura nos testes e 10% menos acurácia durante o treino.

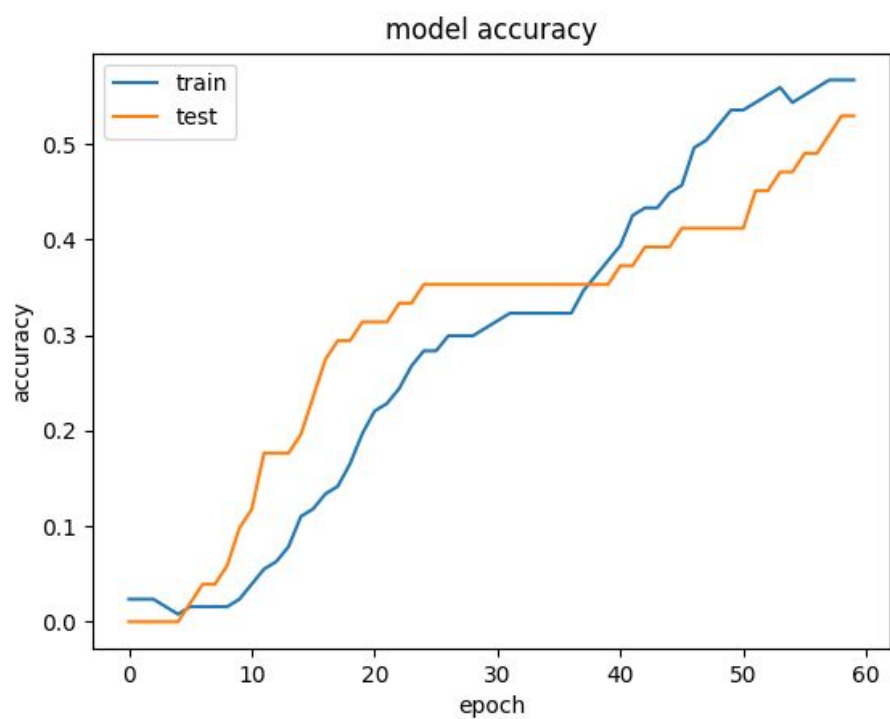


Figura 1: Acurácia do MLP + PCA em função do número de épocas de treinamento

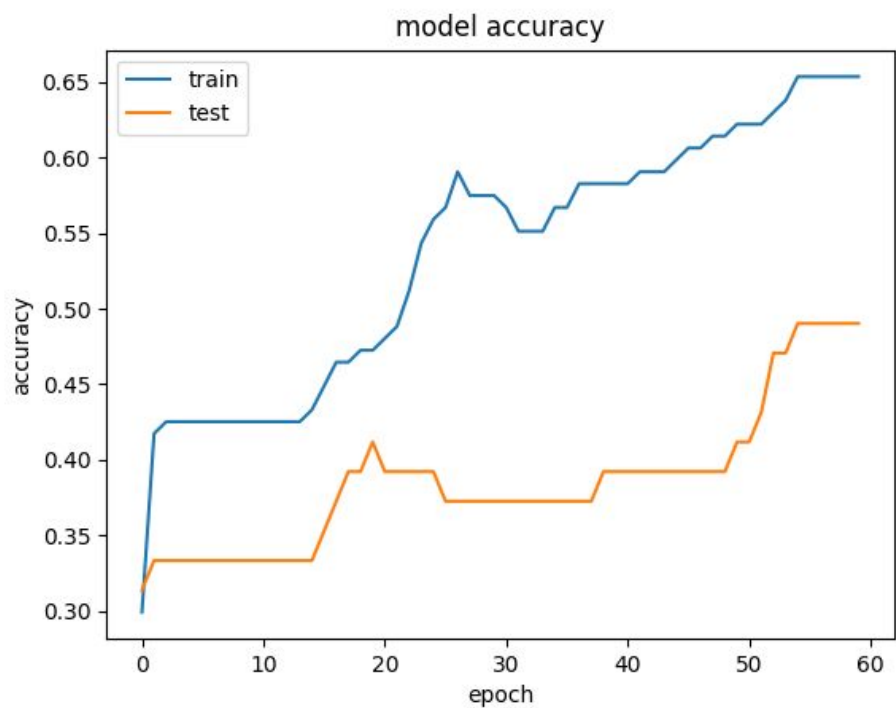


Figura 2: Acurácia do MLP em função do número de épocas de treinamento