

Relatório de Atividades – Projeto III

Neste trabalho avalia-se a eficiência do algoritmo PCA para a redução de dimensionalidade de dados e para a classificação da base de dados *wine*. Foram codificados dois algoritmos. O primeiro é um PCA adaptativo, que é baseado em uma rede neural de Rubner-Tavan e os pesos são atualizados seguindo a regra de Oja. O segundo é um PCA normal, calculado a partir dos autovalores e autovetores retirados da matriz de covariância dos dados.

Os algoritmos PCA são comparados a partir do *scatter plot* das componentes. A eficiência da classificação é comparada a partir da acurácia de uma MLP. A MLP é treinada de três formas: 1) sem utilizar PCA; 2) utilizando o PCA clássico; e 3) utilizando o PCA adaptativo.

Cálculo do PCA

Tanto para o PCA clássico quanto para o PCA adaptativo, os dados são padronizados antes das componentes serem calculadas. Ou seja, é garantido que a média dos dados é igual a zero e a variância é igual a um. Como dito anteriormente, o PCA clássico é calculado a partir dos autovetores.

Por outro lado, o PCA adaptativo é calculado a partir da rede de Rubner-Tavan, que aprende a decorrelação dos dados. Os pesos principais são atualizados seguindo a regra de Oja, e os pesos laterais são computados a partir da regra anti-Hebbiana, que tem por objetivo reduzir a correlação das unidades de saída da rede.

Tendo como base o exercício 5, é sabido que a base de dados *wine* pode ser representada em 3 componentes. Para ambos os modelos de PCA, utiliza-se 3 componentes para fazer o treinamento e teste da MLP e 2 componentes para plotar os resultados.

MLP

A MLP é codificada utilizando os *frameworks* *Tensorflow* e *Keras*. Ela recebe como entrada um vetor 178 x 13 ou 178 x 3. No primeiro caso os dados não foram redimensionados. No segundo caso foi aplicado ou o PCA clássico ou o PCA adaptativo. Foram usados 20% dos dados para teste e 80% para treino. Para todos os casos a rede foi treinada com 2000 épocas.

A rede possui uma camada escondida com 60 neurônios, e uma camada *softmax* com 3. A função de ativação usada é a ReLU e a métrica de avaliação é a acurácia. Escolheu-se uma rede neural com uma camada para avaliar melhor a acurácia da rede. Como a base de dados *wine* não possui uma classificação linear, espera-se que a MLP não tenha uma acurácia muito alta com apenas uma camada e que ao utilizar o PCA essa acurácia melhore.

Organização do Código

O código está dividido em três arquivos: 1) *NNData*, que possui funções para ler a base de dados e fazer uma divisão estratificada dela para treino e teste; 2) *utils*, que possui as funções que calculam os PCAs clássico e

adaptativo, plotam os resultados e rodam a MLP; e 3) *runanalysis*, que roda todas as funções necessárias para analisar a eficácia do PCA como requisito.

Resultados e Conclusão

As Figuras 1 e 2 mostram o *scatter plot* das 2 principais componentes da base de dados *wine* para o PCA clássico e adaptativo respectivamente.

Pode-se observar que para ambos casos todas as classes foram agrupadas de maneira similar, porém o PCA adaptativo aparenta ser uma rotação do PCA clássico em torno dos eixos X e Y, onde a classe 1 (verde) seria o centro da rotação.

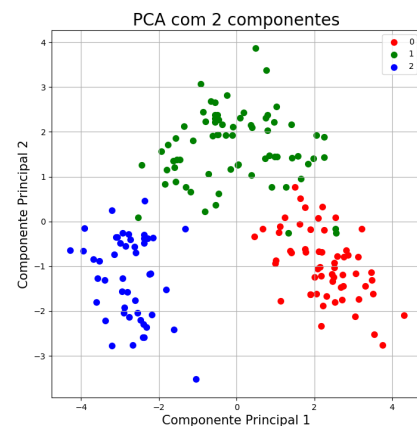


Figura 1: PCA Clássico

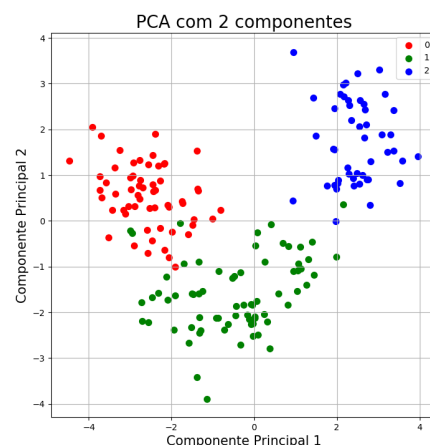


Figura 2: PCA Adaptativo

As Figuras 3, 4 e 5 mostram a acurácia e o erro da classificação durante o treinamento para a MLP sem PCA, com PCA clássico e com PCA adaptativo.

Pode-se observar que no caso da MLP sem PCA a acurácia se manteve próxima de 90%, e para ambos os casos com PCA ela se manteve próxima de 100%. Como esperado, o PCA melhorou a acurácia da classificação. No que diz respeito a diferença entre os PCAs clássico e adaptativo,

aparentemente o clássico tem um aprendizado mais suave que o adaptativo. Além disso, inicialmente o PCA adaptativo possui um erro muito maior que o PCA clássico, porém após poucas épocas ambos convergem para 0%.

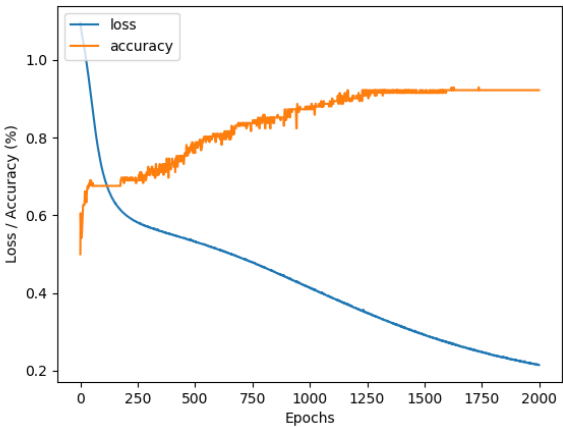


Figura 3: Aprendizado da MLP sem PCA

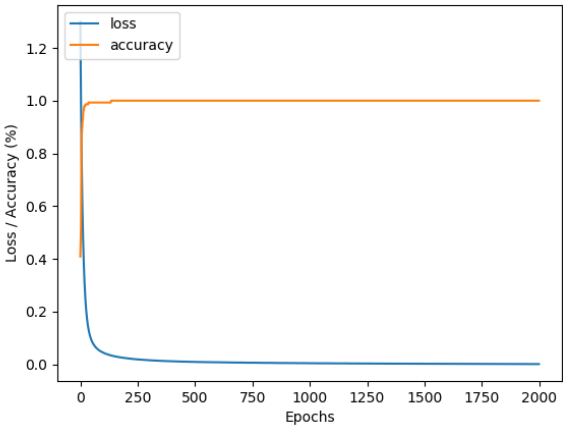


Figura 4: Aprendizado da MLP com PCA clássico

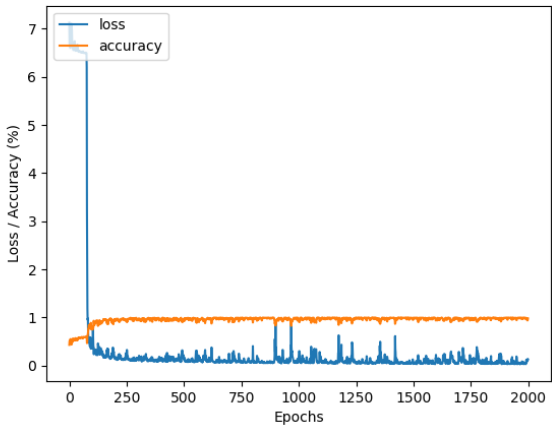


Figura 5: Aprendizado da MLP com PCA adaptativo

A Tabela 1 mostra a acurácia da classificação nos conjuntos de teste. Pode-se observar que a acurácia da PCA adaptativa é melhor que a MLP sem PCA, como esperado, e a MLP sem PCA é melhor que a PCA clássica, o que foi uma surpresa. Acredita-se que pelo fato da divisão da base de dados em conjuntos de treino e teste ser aleatória para cada caso estudado, a seleção das amostras de teste pode estar influenciando o resultado. Vale salientar que durante o treinamento a rede não teve contato com os dados de teste.

	Acurácia	Erro
MLP sem PCA	0.91	0.21
MLP PCA Class.	0.88	1.34
MLP PCA Adapt.	0.94	0.16

Tabela 1: Resultado nos conjuntos de teste

Conclui-se então que o PCA consegue separar bem as componentes principais da base de dados tanto para o caso do PCA clássico quanto para o PCA adaptativo, contudo um é a rotação do outro.

Por fim, esperava-se que a MLP não conseguisse classificar bem a base de dados sem PCA pois a base *wine* não possui uma classificação completamente linear. Porém, como foram utilizados 60 neurônios na camada escondida a rede aprendeu uma função de classificação suficientemente satisfatória. Talvez com menos neurônios a classificação piore consideravelmente.