

Deep Learning - Relatório 5:

Deep K-Means

Rian Freitas e Hanna Rodrigues

Novembro 2023

1 Introdução

Neste trabalho, investigamos e comparamos o desempenho de diferentes configurações do Deep K-Means na tarefa de clusterização conjunta e na síntese dos dados usando o conjunto de dados de dígitos MNIST. Este consiste-se em um conjunto de imagens da escala cinza de dígitos manuscritos. Utilizaremos duas variantes do modelo, o DKMa que usa estratégia de anelamento no alfa, e o DKMp, que é inicializado com um encoder pré-treinado.

2 Variando o parâmetro alfa

A tabela a seguir apresenta o desempenho do Deep K-Means com diferentes configurações do parâmetro alfa na tarefa de clusterização conjunta utilizando o conjunto de dados MNIST. As configurações variam entre o uso da estratégia de anelamento no alfa e a inicialização com um encoder pré-treinado. Foi utilizado 10 como numero de clusters.

| Configuration | α -Pretrained | α -Annealing | Accuracy |
|---------------|----------------------|---------------------|----------|
| 1 | | $\alpha_0=0.01$ | 0.39377 |
| 2 | | $\alpha_0=0.1$ | 0.73666 |
| 3 | | $\alpha_0=0.5$ | 0.30717 |
| 4 | | $\alpha_0=1$ | 0.20893 |
| 5 | $\alpha_n=3$ | | 0.36127 |
| 6 | $\alpha_n=30$ | | 0.83896 |
| 7 | $\alpha_n=150$ | | 0.85026 |
| 8 | $\alpha_n=300$ | | 0.84900 |

Ao analisar as configurações de 1 a 4, onde a estratégia de anelamento é empregada, é evidente uma variação significativa na acurácia conforme os diferentes valores de alfa zero. Na configuração 1, observa-se uma dispersão considerável nos dados por meio da matriz de confusão, sugerindo que possivelmente o valor de alfa está muito baixo, tornando a atribuição excessivamente flexível. Essa

hipótese é corroborada pelos resultados promissores do experimento 2, onde houve um aumento substancial na acurácia do modelo e uma maior concentração na matriz de confusão.

Entretanto, persistem confusões nas imagens dos dígitos "4" e "9", indicando que seus clusters estão próximos entre si em comparação com os outros dígitos. Nas configurações 3 e 4, de modo análogo, o aumento do alfa resultou em uma atribuição rígida, com o modelo atribuindo apenas às classes "0", "1", "3" e "9", consequentemente resultando em uma acurácia inferior ao experimento 2.

Concluimos, portanto, que o anelamento não deve ser excessivamente restritivo a ponto de inibir a atribuição de classes, nem excessivamente flexível a ponto de permitir uma grande dispersão na distribuição das classes.

Nas configurações de 5 a 8, onde a inicialização do modelo é realizada com um encoder pré-treinado, observa-se consistentemente acurácias mais elevadas. Isso sugere que a transferência de conhecimento do encoder pré-treinado é benéfica para o desempenho do modelo na tarefa de clusterização. Configurações com valores mais elevados de α , como 30, 150 e 300, apresentam uma tendência crescente na acurácia.

Contudo, um desafio semelhante aos experimentos com anelamento persiste: há confusões nas imagens dos dígitos "4" e "9". Isso reforça a hipótese de que no espaço latente, essas duas classes estão muito próximas uma da outra, contribuindo para as dificuldades do modelo em distingui-las com precisão.

3 Variando o número de clusters

Na segunda etapa deste estudo, exploramos a influência do número de clusters na melhor configuração do modelo, inicialmente otimizada para 10 clusters (Configuração 7). A variação foi realizada considerando os números de clusters igual a 2, 5, 10 e 20.

Ao utilizar 2 clusters, observamos uma simplificação na distinção entre os dígitos, o que potencialmente gera confusão entre classes, dado que apenas duas classes estão disponíveis para atribuição. No entanto, a observação de que "9" e "6" não compartilham o mesmo cluster sugere que a distância entre esses dígitos no espaço latente é maior do que a distância entre "6" e "0", que compartilham o mesmo cluster.

Ao expandir o número de clusters para 5, notamos a proximidade dos dígitos "9", "7" e "4", pertencendo ao mesmo cluster. Contrariamente, o dígito "2" se destaca dos demais, possivelmente devido à sua maior distância no espaço latente.

Contudo, é importante ressaltar que um aumento excessivo no número de clusters pode resultar em segmentações muito específicas, prejudicando a capacidade de generalização do modelo. Isso é evidenciado quando o número de clusters atinge 20, ocasionando uma piora na acurácia, indicando que a atribuição está segregando em clusters diferentes dados que já estão muito próximos. Encontrar um equilíbrio adequado no número de clusters é crucial para evitar uma segmentação excessiva e garantir que o modelo capture padrões representativos

dos dados.

Aqui, segue o [\[link do Colab\]](#).

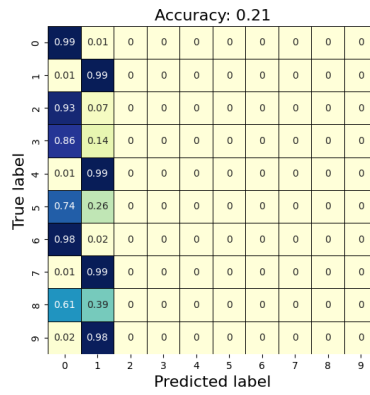


Figure 1: Matriz de confusão para 2 clusters

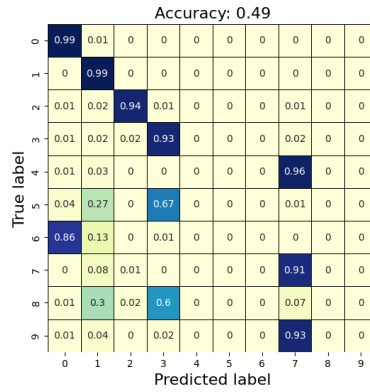


Figure 2: Matriz de confusão para 5 clusters

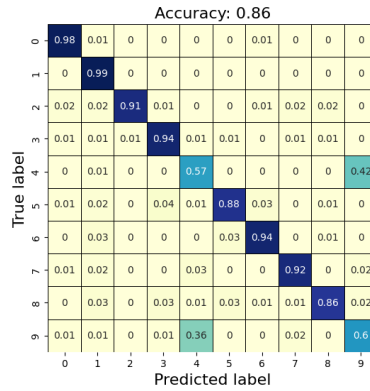


Figure 3: Matriz de confusão para 10 clusters

