

# Comparação de desempenho entre implementações do K-Means

João Augusto Dias Neto, Rafael Rodrigues de Souza

## 1 Introdução

Este relatório tem como objetivo comparar a eficiência e a precisão de duas implementações do K-Means: uma versão manual desenvolvida do zero e a versão disponível na biblioteca *sklearn*. Para medida de comparação, os testes em ambas as implementações foram avaliados utilizando a mesma base de dados e diferentes abordagens, incluindo os dados originais e reduções de dimensionalidade via PCA (2D e 1D).

## 2 Dados Coletados

### 2.1 $k = 3$

Abordagem	Iterações	Silhouette Score	Tempo (s)
Original (Manual)	6	0.5528	0.0028
PCA 2D (Manual)	7	0.5977	0.0025
PCA 1D (Manual)	8	0.6771	0.0043
Original ( <i>sklearn</i> )	7	0.5512	0.0079
PCA 2D ( <i>sklearn</i> )	8	0.5977	0.0278
PCA 1D ( <i>sklearn</i> )	11	0.6771	0.0154

Table 1: Comparação de iterações, Silhouette Score, tempo e uso de memória para diferentes abordagens entre a implementação manual e a do *sklearn* para  $k = 3$ .

### 2.2 $k = 5$

Abordagem	Iterações	Silhouette Score	Tempo (s)
Original (Manual)	5	0.4931	0.0016
PCA 2D (Manual)	7	0.5510	0.0013
PCA 1D (Manual)	8	0.6815	0.0013
Original ( <i>sklearn</i> )	7	0.4610	0.0111
PCA 2D ( <i>sklearn</i> )	6	0.5460	0.0248
PCA 1D ( <i>sklearn</i> )	5	0.6823	0.0162

Table 2: Comparação de iterações, Silhouette Score, tempo e uso de memória para diferentes abordagens entre a implementação manual e a do *sklearn* para  $k = 5$ .

Memória Inicial: 211.32 MB — Memória Final(Manual): 215.07 MB — Memória Final(Sklearn): 211.32 MB

## 3 Discussão dos Resultados

A implementação manual foi mais eficiente em tempo e iterações, especialmente para  $k=3$ . A versão *sklearn* foi mais estável em termos de memória, mantendo 211.32 MB, enquanto a manual utilizou 215.07 MB. Ambos os métodos produziram resultados semelhantes, com a manual sendo superior em termos de performance.

## 4 Conclusão

Ambas as implementações foram adequadas, com a versão manual sendo mais rápida, enquanto a *sklearn* apresentou maior estabilidade. A redução de dimensionalidade com PCA melhorou o Silhouette Score, especialmente no caso 1D.