

K-Means – דו"ח

פרטים כלליים על האלגוריתם ומימוש:

אתחול הסנטרואידים – נעשה ע"י שימוש בפונקציה הנתונה `init_centroids`.
 איטרציות – עבור כל K (2,4,8,16) נבצע 10 איטרציות שבסופן תוצג התמונה המכווצת המורכבת מ- K צבעים.
 בכל אחת מ-10 האיטרציות מתבצעים שני שלבים:

1. התאמה בין פיקסל לבין סנטרואיד – נשייך כל פיקסל בתמונה לסנטרואיד הקרוב ביותר אליו (נעשה זאת על ידי חישוב מרחק אוקלידי בין פיקסל לכלל הסנטרואידים ונבחר את הסנטרואיד עם המרחק המינימלי).
2. עדכון הסנטרואידים – נחשב את המרחק הממוצע בין סנטרואיד מסוים לכל הפיקסלים המשוייכים אליו, ונעדכן את ערך הסנטרואיד להיות ממוצע זה.

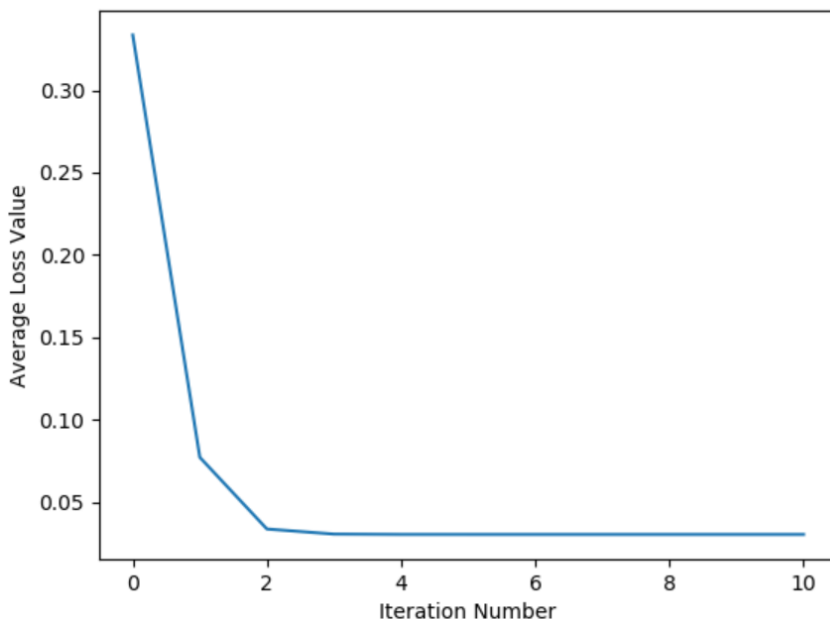
הדפסת הפלט- נעשה ע"י שימוש בפונקציה הנתונה `print_cent` (הצגת התמונה עצמה ע"י שימוש בספריית `matplotlib`).

לכל K (בטווח הערכים המבוקש- 2,4,8,16) מוצגת התמונה המכווצת שמתקבלת, וכן גרף של פונקציית ה-Loss. ניתן לראות בבירור קשר ישיר בין כמות האיטרציות למדד השגיאות- ככל שמתבצעות יותר איטרציות כך האלגוריתם לומד עד שמתכנס לנקודה סופית לכל סנטרואיד כך שה-Loss מינימלי- כלומר מזערנו את ממוצע השגיאות כנדרש, וכן סיווגנו את הנקודות בצורה נכונה.

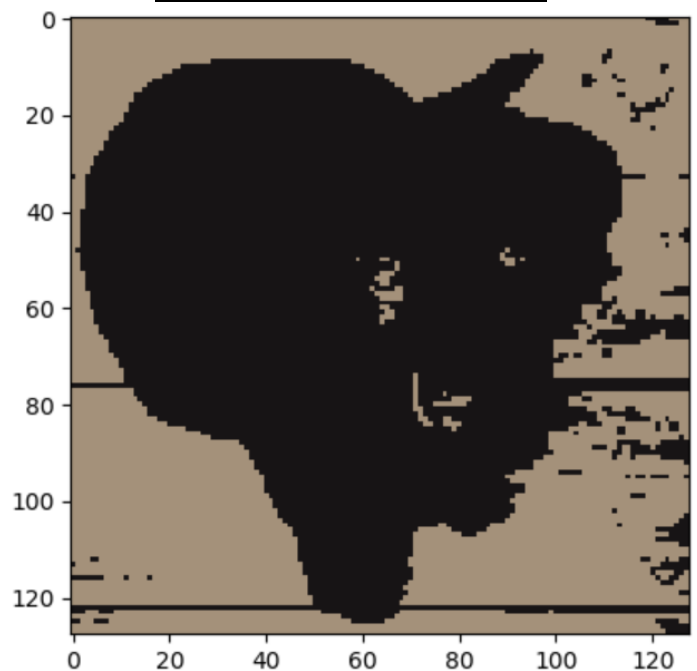
עבור $K = 2$:

Average Loss value

$k = 2$:



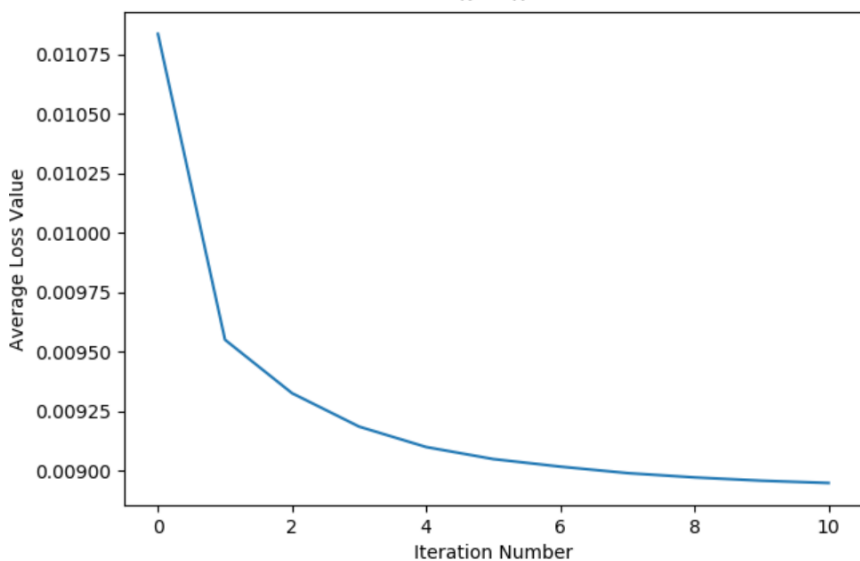
Compressed with 2 colors



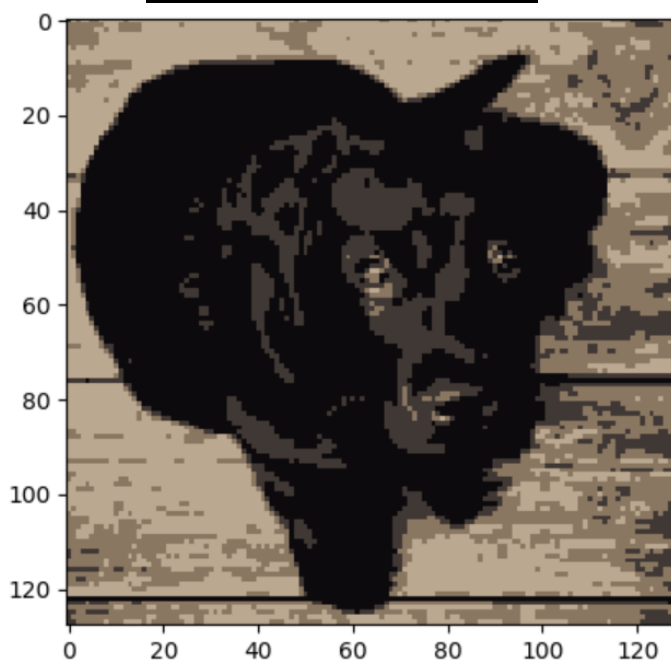
עבור $K = 4$:

Average Loss value

$k = 4$:



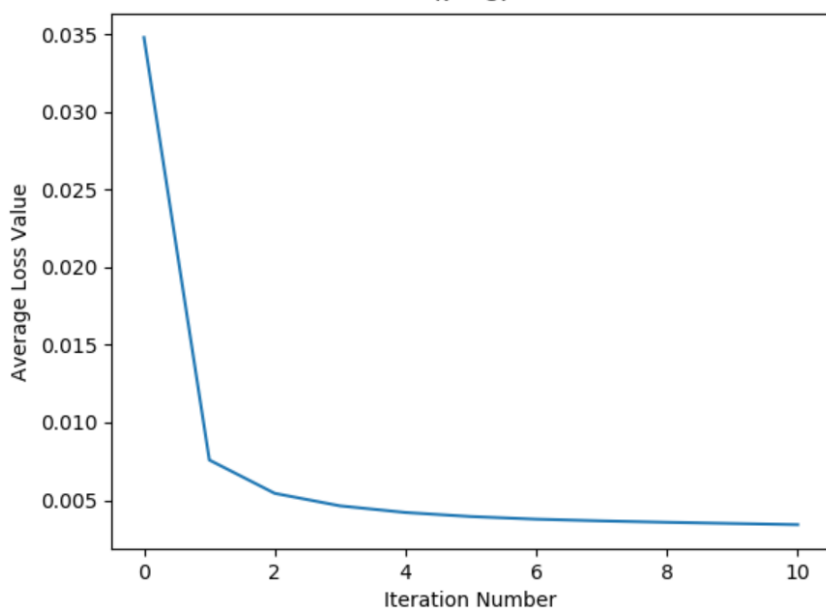
Compressed with 4 colors



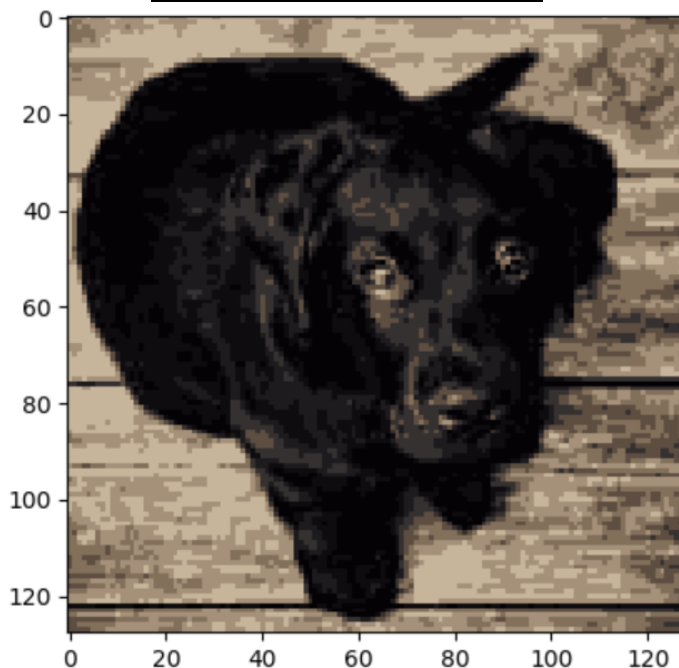
עבור $K = 8$:

Average Loss value

$k = 8$:



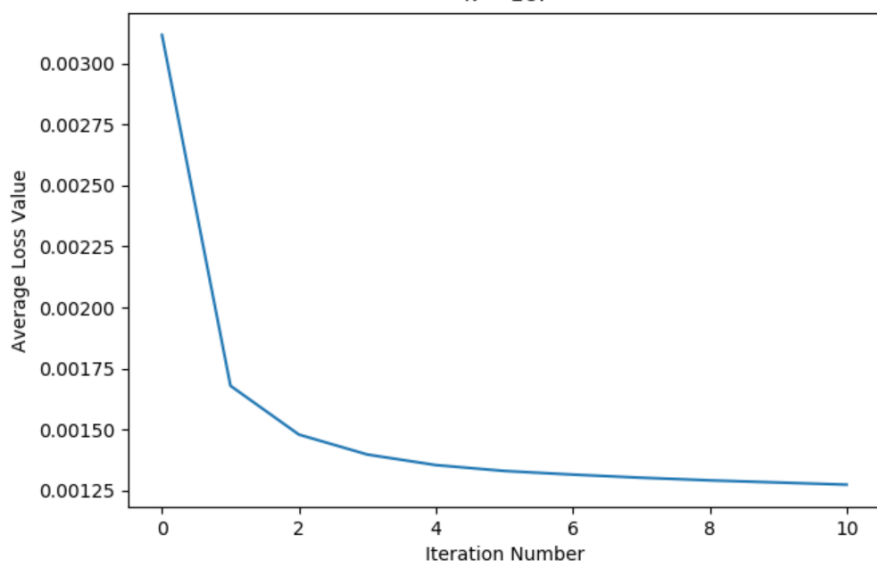
Compressed with 8 colors



עבור $k = 16$:

Average Loss value

$k = 16$:



Compressed with 16 colors

