# MACHINE LEARNING

# <u>Assignment 5 – K-Means Clustering</u>

## 1. Koala:

Original Image





# • <u>K</u> = 5



# • <u>K = 10</u>



### • K = 15



#### • K = 20



• It is clear from the observation of the above images that the number of colors in the compressed imaged depends on the 'k' value used for the K-Means algorithm.

- We can see a drastic difference in the quality of the image when we increase the k value from 5 to 10. But still, it is missing a few more details, especially on the left side of the image. K = 15 gives a more detailed image compared to k = 10.
- Although k = 20 can give a better image than k = 15, That much increase in complexity is not worth the improvement in quality. So, I feel the best k value for this image is k = 15

### 2. Penguins:

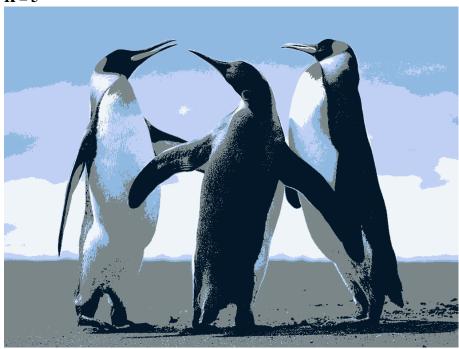
• Original Image



• **K** = 2



# • K = 5



## • K = 10



### • K = 15



### • K = 20



• A similar observation has been seen in the penguin image as the koala bear image. Quality increases with k value.

- Unlike the koala bear image, the compressed image with k = 5 even captures the shadow of the penguins, and k = 10 capture the yellow shade in the face and the body of the penguins.
- K = 15 gives more depth and clarity to the image, and the only difference in k = 20 compared to the previous version is the detailing in the background. But I feel that k = 10 gives a pretty good idea of the image for this case.
- One common observation from both images is that the blurry parts of the original image are hard to capture with lower k values.
- Running K-Means till convergence would take much time, I have run 10 iterations of K-Means for both the images