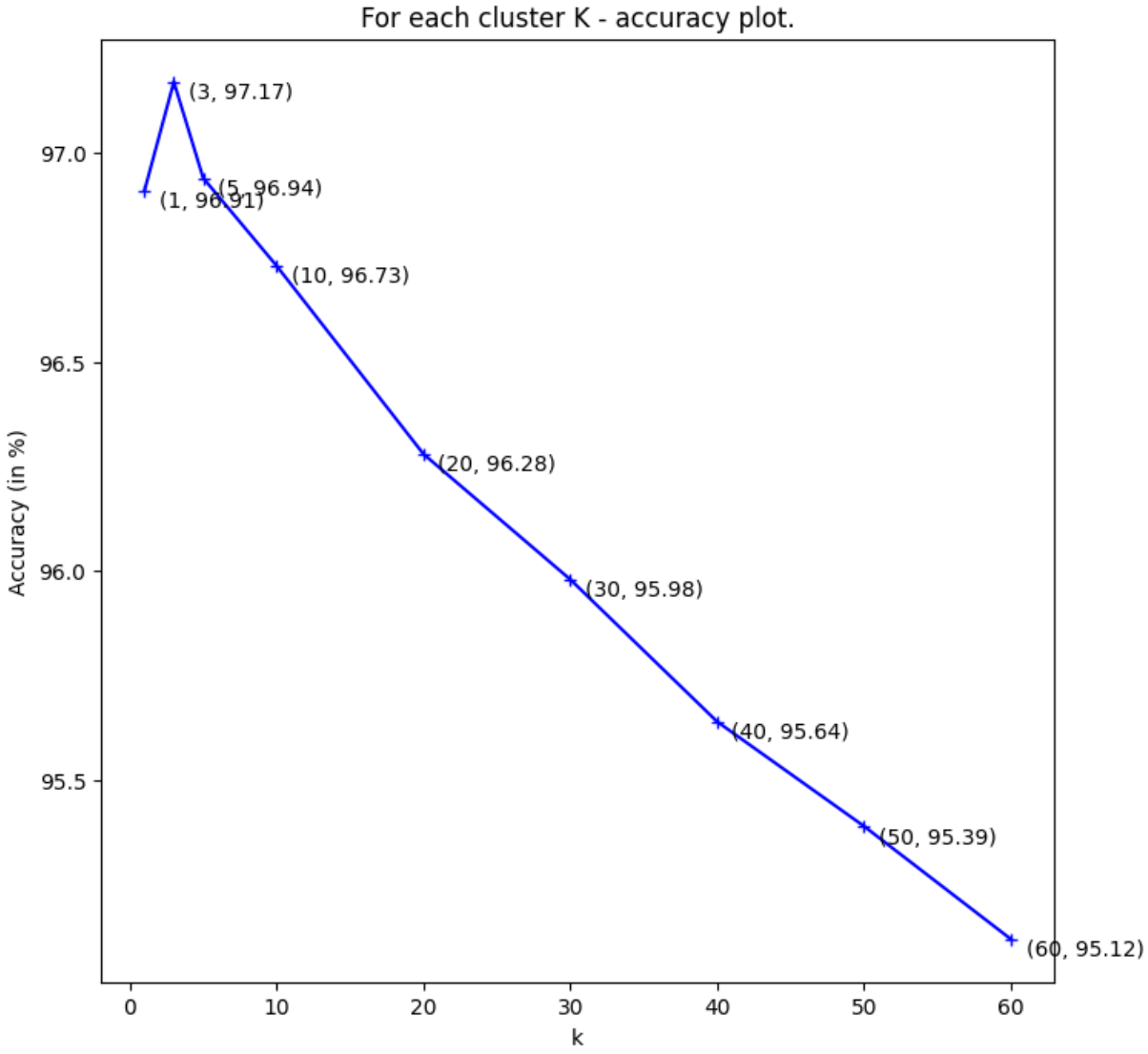
ANS for Q1 - Handwritten Digits Recognition with k-NN

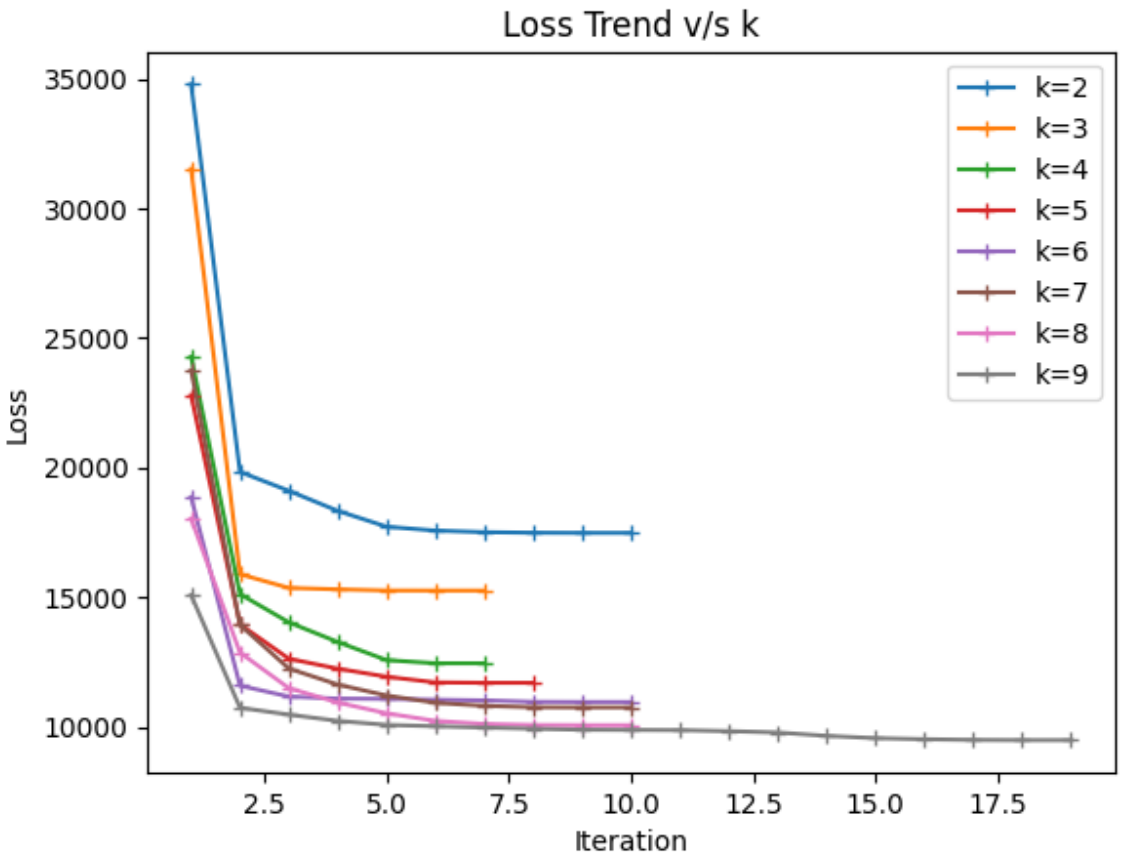


Observations and Explanations:

1. **Accuracy vs. K Plot:**
   - **a.** The accuracy increases from K=1 to K=3, reaching a peak at K=3.
   - **b.** After K=3, the accuracy starts to decrease, indicating that the model might be overfitting to the training data as K increases.
2. **Accuracy Values:**
   - **a.** {1: 96.91, 3: 97.17, 5: 96.94, 10: 96.73, 20: 96.28, 30: 95.98, 40: 95.64, 50: 95.39, 60: 95.12}
3. **Optimal K:**
   - **a.** The optimal K for this dataset appears to be around 3, as it provides the highest accuracy.
   - **b.** A smaller K helps reduce the impact of noise, leading to improved accuracy.

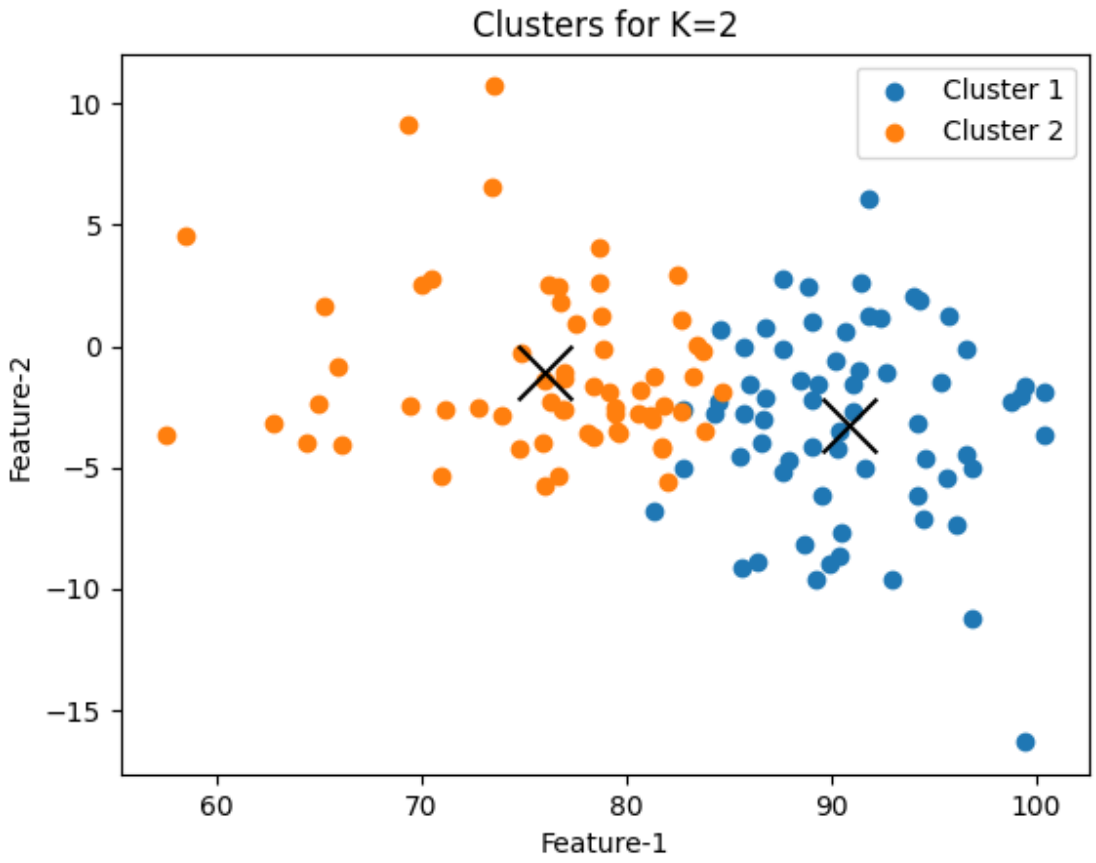
4. **Performance Considerations:**
    a. Precomputation of pairwise distances enhances the efficiency of the k-NN algorithm, making it more scalable for larger datasets.
5. **Dataset Size and Characteristics:**
    a. The dataset size and characteristics play a significant role in determining the optimal K.
    b. Further analysis may involve exploring different distance metrics, dataset preprocessing techniques, and feature representations.
    c. techniques like deskewing, noise removal, blurring will increase the accuracy.

ANS for Q2 - K-Means Clustering



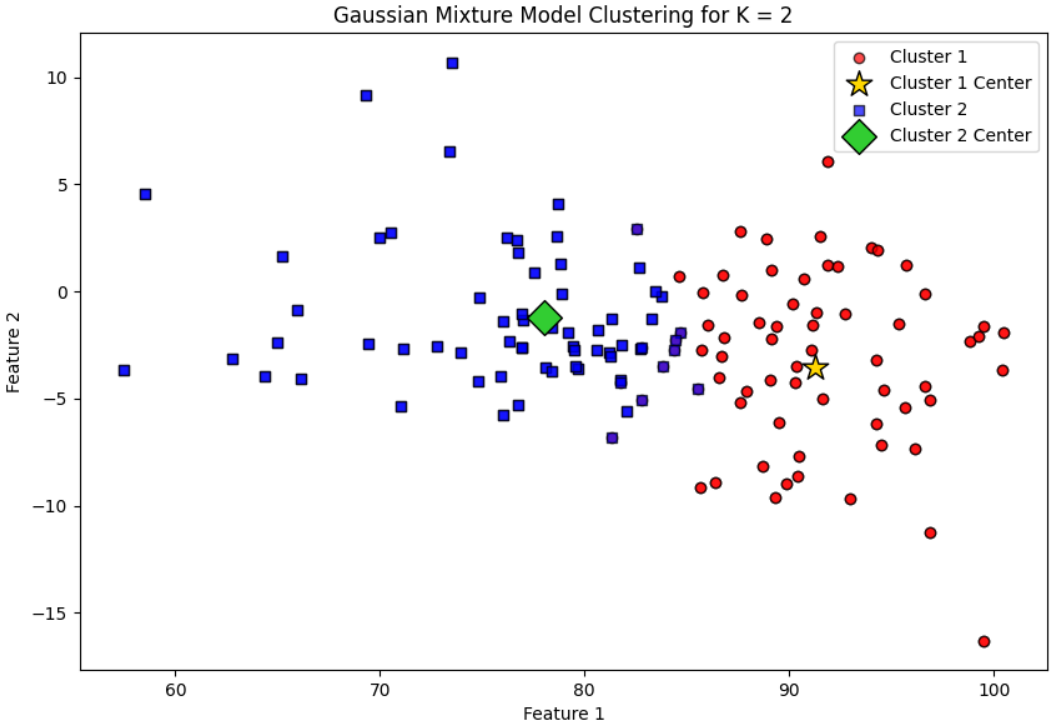
1. **Objective Function (Loss) Trend:**
    a. As K increases, the loss generally decreases. This is expected, as a higher value of K allows the algorithm to create more clusters, potentially better fitting the data points to their respective centroids.
    b. However, it's important to observe the rate of decrease. After a certain point, the reduction in loss becomes less significant with each increase in K. This is consistent with the diminishing returns associated with increasing complexity.

2. **Trade-off and Model Interpretability:**
    a. Choosing the optimal K involves a trade-off between model simplicity and interpretability. A model with too many clusters might capture noise or idiosyncrasies in the data, making it less interpretable and generalizable to new data.
    b.



3. **Visualization for K=2:**
    a. For K=2, the data points are plotted using their first two features. Two different colors or symbols distinguish the two clusters. This visualization helps in understanding how the algorithm partitions the data into distinct groups.
    b. If the two clusters are well-separated and visually distinct, it suggests a reasonable partitioning. However, if there is significant overlap, it might indicate that two clusters are not sufficient.

ANS for Q3 - Gaussian Mixture Model



1. **Convergence Criteria:** The implementation follows the recommendation to terminate the iteration when the change in log-likelihood falls below a small threshold (e.g., 1e-5). This ensures that the Gaussian Mixture Model (GMM) reaches a stable state.
2. **Similarity to K-Means Plot:** Plot of the GMM is very similar to the plot for K-Means, indicating comparable clustering results in this specific case with K=2.
3. **Color-Coding of Clusters:** The clusters are represented in red for Feature 1 and blue for Feature 2. This color-coding provides a clear distinction between the clusters and aids in the visual interpretation of the results.

THE END