Clustering-4

Assignment Questions

PW SKILLS

Q1. Explain the concept of homogeneity and completeness in clustering evaluation. How are they calculated?

Q2. What is the V-measure in clustering evaluation? How is it related to homogeneity and completeness?

Q3. How is the Silhouette Coefficient used to evaluate the quality of a clustering result? What is the range of its values?

Q4. How is the Davies-Bouldin Index used to evaluate the quality of a clustering result? What is the range of its values?

Q5. Can a clustering result have a high homogeneity but low completeness? Explain with an example.

Q6. How can the V-measure be used to determine the optimal number of clusters in a clustering algorithm?

Q7. What are some advantages and disadvantages of using the Silhouette Coefficient to evaluate a clustering result?

Q8. What are some limitations of the Davies-Bouldin Index as a clustering evaluation metric? How can they be overcome?

Q9. What is the relationship between homogeneity, completeness, and the V-measure? Can they have different values for the same clustering result?

Q10. How can the Silhouette Coefficient be used to compare the quality of different clustering algorithms on the same dataset? What are some potential issues to watch out for?

Q11. How does the Davies-Bouldin Index measure the separation and compactness of clusters? What are some assumptions it makes about the data and the clusters?

Q12. Can the Silhouette Coefficient be used to evaluate hierarchical clustering algorithms? If so, how?

Note: *Create your assignment in Jupyter notebook and upload it to GitHub & share that github repository link through your dashboard. Make sure the repository is public.*