

بسمه تعالی

سری دوم سوالات مباحث ویژه در داده کاوی

داده های ORL را در نظر بگیرید

- ۱- از روشهای مختلف خوشه بندی مانند کا- میانگین - DBSCAN و روش مبتنی بر چگالی برای خوشه بندی این دادهها استفاده نمایید. با چند تا از روشهای ارزیابی مانند NMI آنها را ارزیابی نمایید.
- ۲- الف) با استفاده از دو روش ساده گفته شده در کلاس بر اساس یک شعاع و همچنین کا- همسایگی گراف متناظر با این داده ها را بدست آورید:
- در اینجا سیگما و دلتا را برابر با میانگین توان دوم فاصله تمام داده ها از همدیگر در نظر بگیرید. همچنین در روش مبتنی بر کا- همسایگی از $k=5$ استفاده نمایید.
- ب) از روش OMP برای ساختن گراف شباهت این داده ها استفاده نمایید.
- ج) از روش خوشه بندی طیفی مبتنی بر Ratio-cut و Normalized-Cut برای خوشه بندی استفاده نمایید.