

Appendix

A. Comparison to other methods

In this section, we compare DBSCAN++ against replacing the nearest-neighbor search needed for DBSCAN with an approximate nearest neighbor method using the FLANN (<https://www.cs.ubc.ca/research/flann/>) library, and we call it ANN DBSCAN.

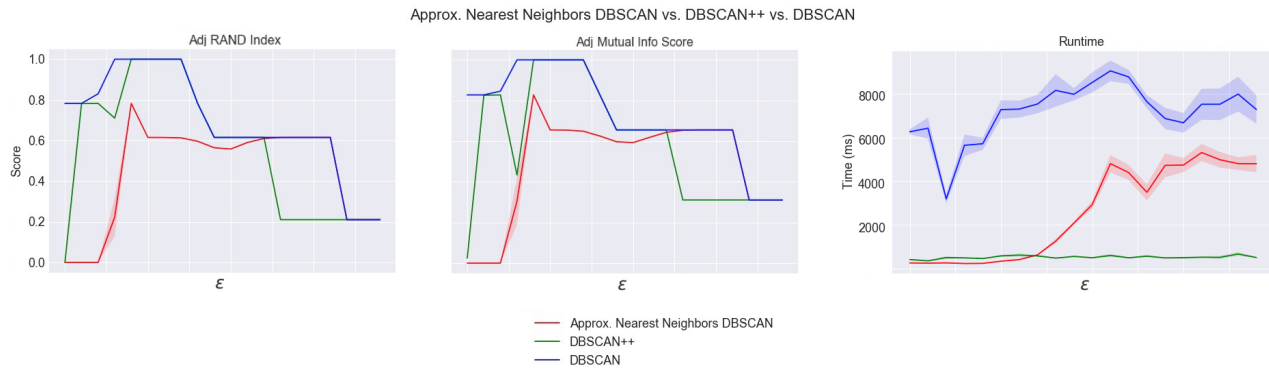


Figure 9. DBSCAN using approximate nearest neighbors vs. DBSCAN++ vs. DBSCAN. Experimental results on a synthetic dataset of 10,000 points drawn from five 50-dimensional uniform distributions run on DBSCAN++, DBSCAN, and DBSCAN using a fast approximate nearest neighbors algorithm from the FLANN library. DBSCAN++ was run with K -center initialization and $m/n = 0.1$. All algorithms were run with $minPts = 10$. ANN DBSCAN shows a comparable speedup to DBSCAN++ but poorer performance compared to both DBSCAN and DBSCAN++, whereas DBSCAN++ shows both comparable performance to DBSCAN and comparable runtime to ANN DBSCAN.