
R W R DARLING, U.S. Department of Defense

Efficient comparison-based learning via partitioned local depth for near neighbors

A triplet comparison on a set S takes $x \in S$, and for any pair $\{y, z\} \subset S \setminus \{x\}$ declares which of y and z is more similar to x . Such triplet comparisons supply an orientation to the line graph of the complete graph on S : $\{x, y\} \rightarrow \{x, z\}$ means y is more similar to x than z is. No metric is involved. Partitioned local depth (PaLD) supplies a non-parametric partitioning of S , under such triplet comparisons, but suffers from a run time cubic in $n := |S|$. We use the Bounded Differences Inequality to analyze the quality of approximation obtained by restricting PaLD to the K -nearest neighbors of each object. Run time is $O(nK^2)$, while error decays exponentially in $-K^2$. The statistics of uniform random orientations of the line graph play a major role. Examples from spatial data at inhomogenous scales are examined.