The optimal algorithm for a given dataset is a complicated choice, and depends on a number of factors:

* number of samples N (i.e. n\_samples) and dimensionality D (i.e. n\_features).
  + Brute force query time grows as O[DN]
  + Ball tree query time grows as approximately O[Dlog⁡(N)]
  + KD tree query time changes D in a way that is difficult to precisely characterize. For small D (less than 20 or so) the cost is approximately O[Dlog⁡(N)], and the KD tree query can be very efficient. For larger D, the cost increases to nearly O[DN], and the overhead due to the tree structure can lead to queries that are slower than brute force.