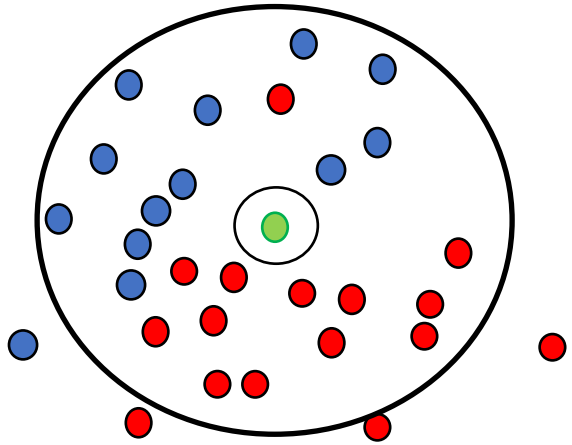
# An adaptive nearest neighbor rule for classification

Akshay Balsubramani, Stanford
Sanjoy Dasgupta, UCSD
Yoav Freund, UCSD
Shay Moran, Google AI Princeton

# Main Idea: Modify *k*-NN Algorithm by Choosing k Adaptively for Each Query

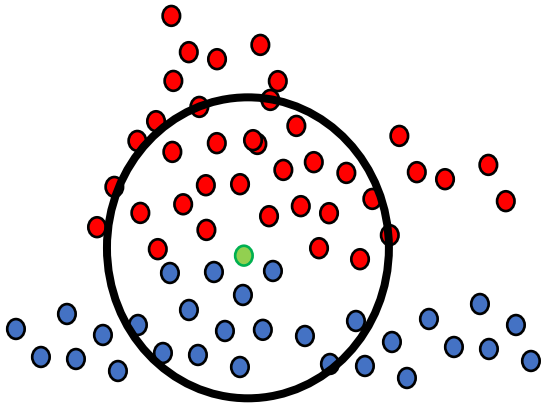
- Classical ***k***-NN: classify *x* by the majority vote of its ***k*** nearest in the training set.



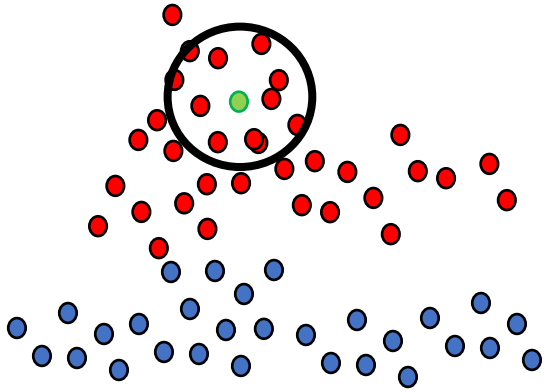
*x* is the green point in the middle. The label assigned to *x* is determined by its ***k*** nearest neighbors (inside the big circle, in this example ***k***=*13*+*12*=*25*)

# Main Idea: Modify ***k***-NN Algorithm by Choosing k Adaptively for Each Query

- Adaptive ***k***-NN:
  - Iterate over the neighbors of *x* from nearest to furthest and query their labels.
  - If one of the label-classes obtains a significant majority then exit the loop and use this label to classify *x*.



Points *x* that are close to the boundary require querying a large number of neighbors



Points *x* that are far from the boundary observe a significant advantage after querying a small number of neighbors

# Main Results

Theoretical Results

1. Adaptive k-NN rule is consistent (i.e. achieves Bayes optimality in the limit).

2. Instance-dependent generalization bounds
   - Number of examples required to classify *x* correctly depends on its "local-margin" (a formal notion introduced in the paper).
   - Points far from the boundary are correctly classified fast.

Practical Results

1. Adaptive *k*-NN rule is competitive with Classical *k*-NN with the **best choice** of *k*

   - Thus, this method circumvents the need to tune the meta-parameter *k.*