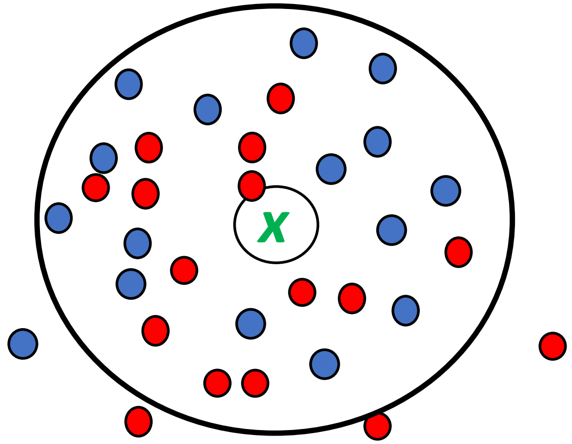
An adaptive KNN Classifier

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

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

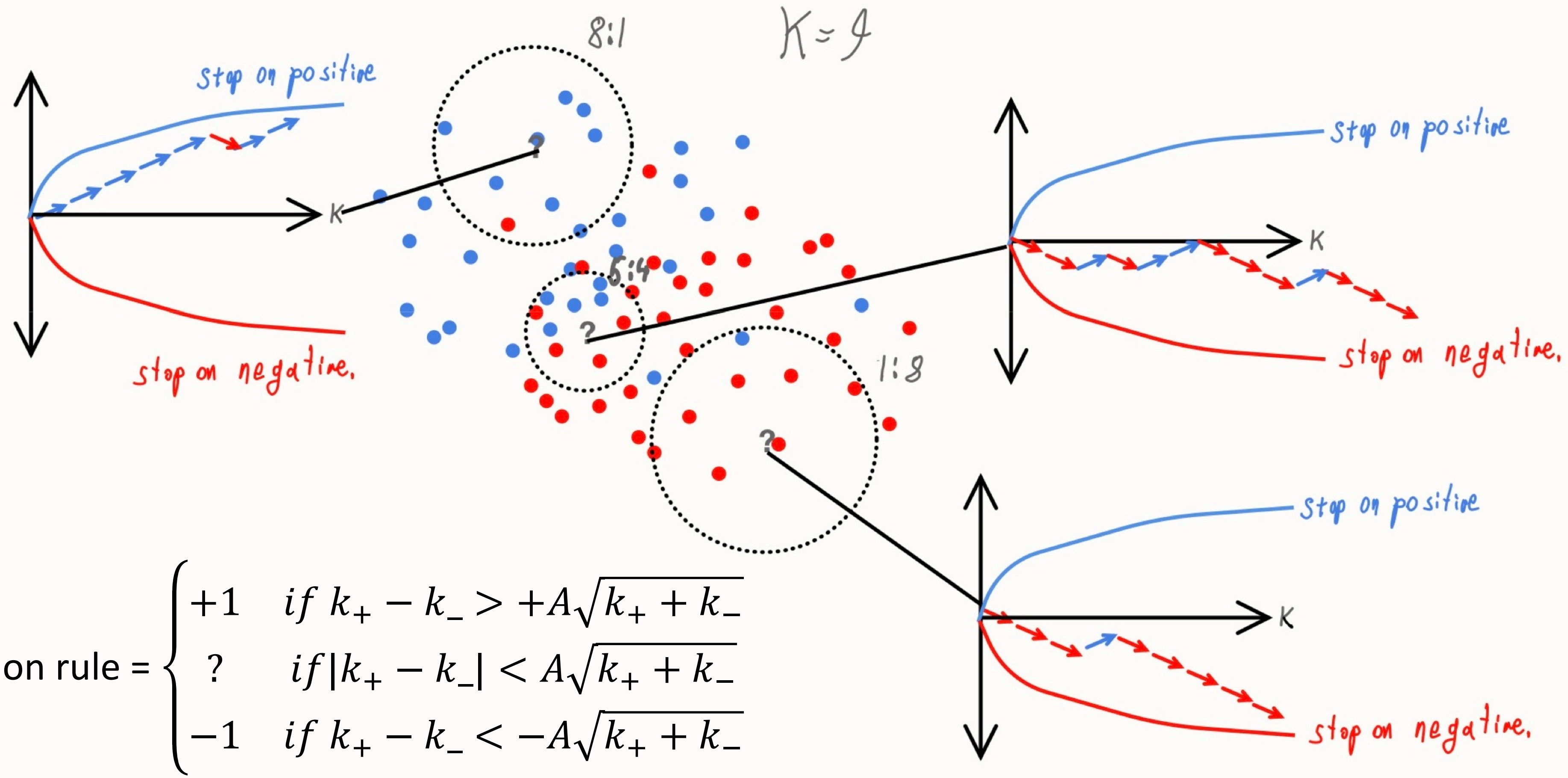


k*=*14*+*15*=*29
14 red
15 blue
Classify as blue

- ***k*NN Converges to Bayes Optimal as $n \to \infty, k \to \infty, \; {}^{k}/_{n} \to 0$**
- How should we choose k for a finite n?
- Different k for different points?
- Should we trust a 15:14 ratio?

Main Idea: Modify ***k***-NN Algorithm by Choosing ***k*** Pointwise

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



Prediction rule = $\begin{cases} +1 & if\ k_+ - k_- > +A\sqrt{k_+ + k_-} \\ ? & if\,|k_+ - k_-| < A\sqrt{k_+ + k_-} \\ -1 & if\ k_+ - k_- < -A\sqrt{k_+ + k_-} \end{cases}$

Theoretical Results

1. Adaptive k-NN rule is consistent (i.e. achieves Bayes optimality in the limit).
2. Adaptive *k*-NN rule is competitive with Classical *k*-NN with the **best choice** of k
3. <u>Pointwise</u> Generalization Bounds
    - Number of examples required to classify *x* correctly depends on its "local-advantage" (a formal notion introduced in the paper).
    - Points far from the boundary are correctly classified fast.

Experimental Results

Not-MNIST