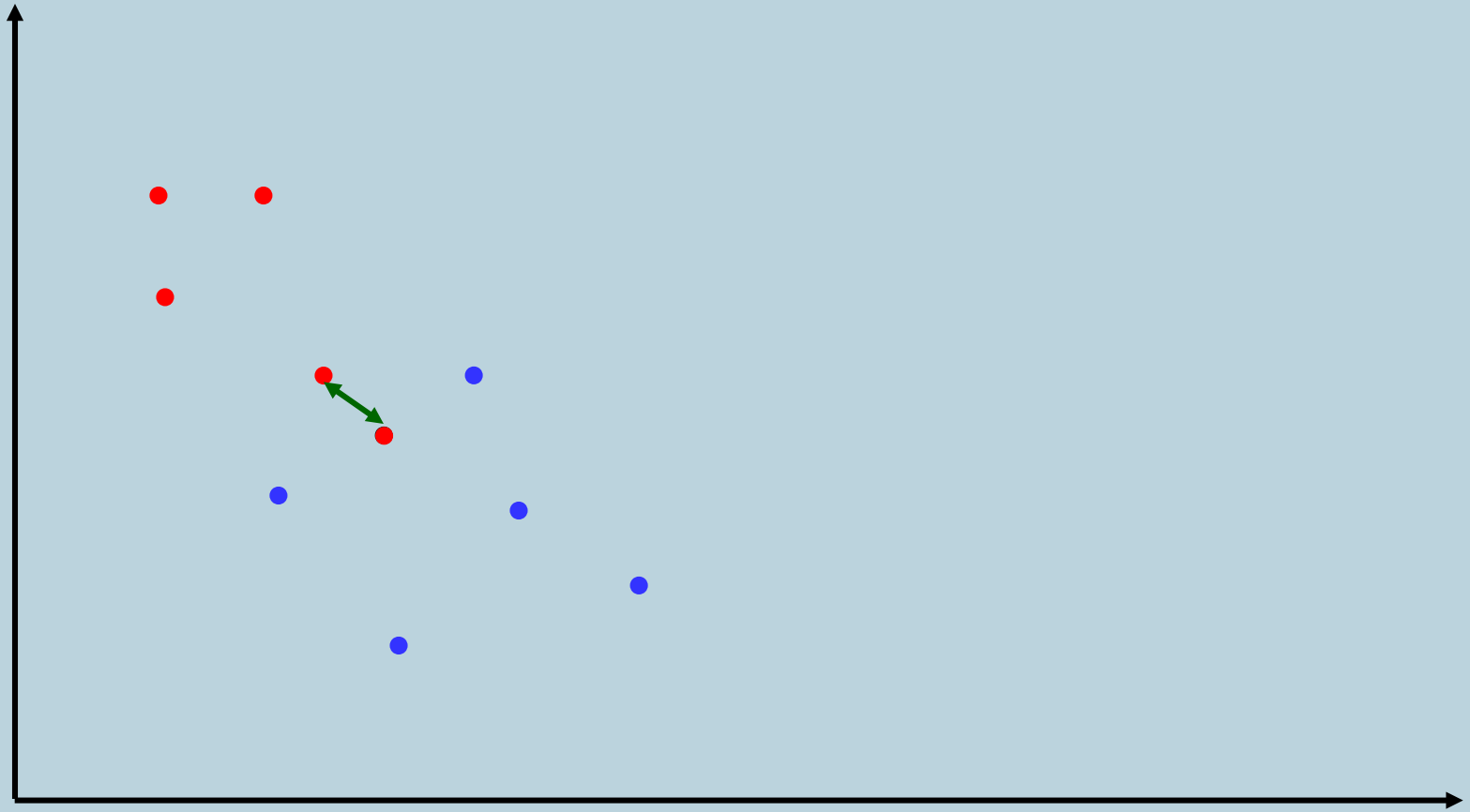
Analogical / Instance-Based

- Avoids learning an abstract concept, but instead reasons from individual cases.
- Simplest approach is "nearest neighbor":
  - During training, simply store the examples, possibly indexing them to support efficient retrieval.
  - For a test case, find the most similar example in the stored training data.
  - Assign the test example to the category of the most similar training instance.

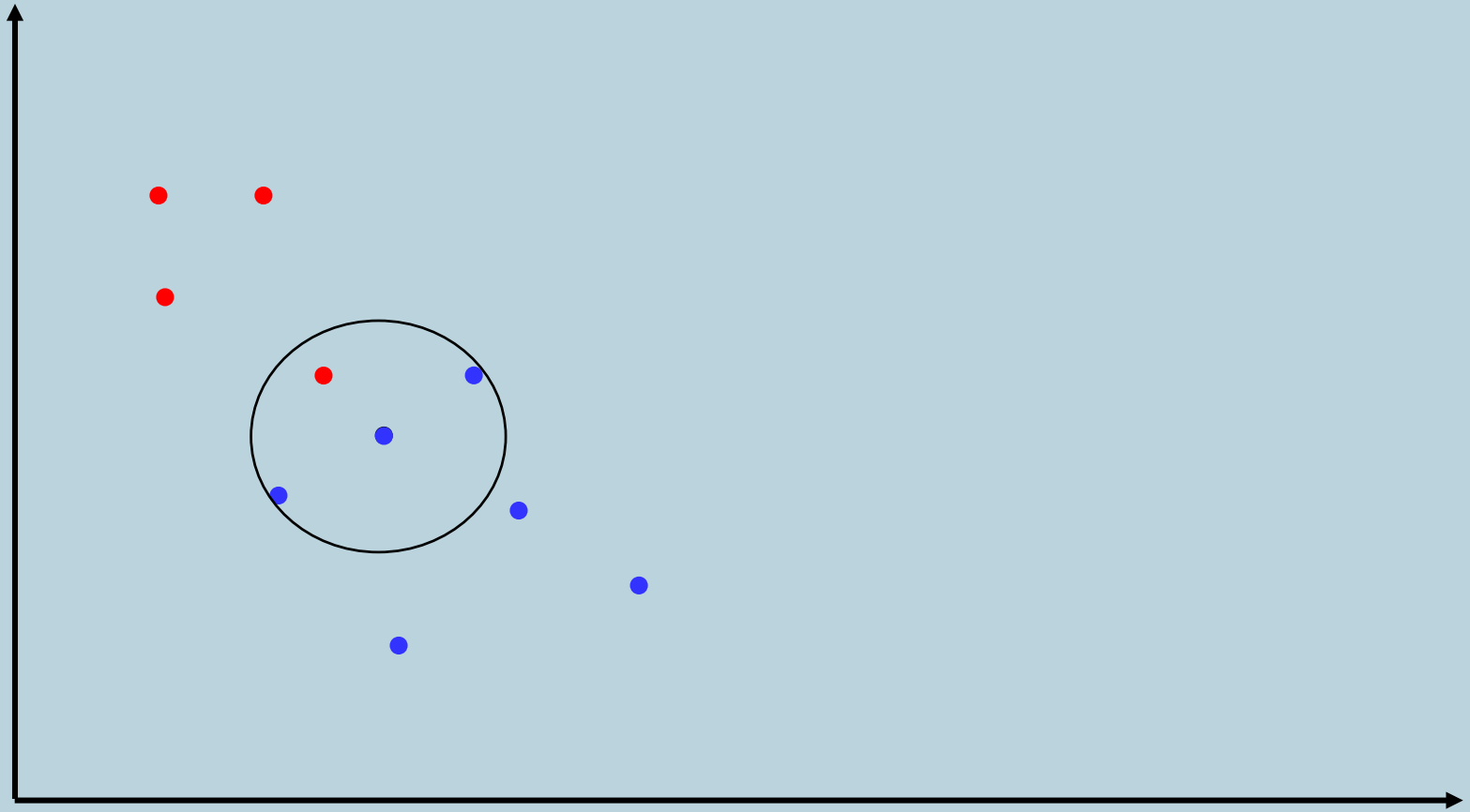
Nearest Neighbor Illustration



K Nearest Neighbor

- Normal nearest neighbor is subject to misclassifying an example due to a single outlier or incorrectly labeled training example.
- “K nearest neighbor” improves robustness by taking the majority vote among the K closest examples.

3 Nearest Neighbor Illustration



Weaknesses of Instance-Based Approach

- Assumes that an appropriate measure of similarity is available.
- Requires effective indexing method to efficiently retrieve nearest neighbors to avoid the costly task of examining all the training examples.