CS194 Fall 2011 Lecture 10
Instance-Based Methods
Choosing \( k \) in k-nearest-neighbors

- \( k=1 \)
- \( k=7 \)
- \( k=15 \)

Training Error: 0.145
Test Error: 0.225
Bayes Error: 0.210
Cross-validation

The graph shows the misclassification errors as a function of the number of neighbors. The lines represent different error measures:
- **Test Error** (blue line)
- **10-fold CV** (green line)
- **Training Error** (orange line)
- **Bayes Error** (purple line)

The y-axis represents the misclassification errors, and the x-axis represents the number of neighbors.
Choosing $k$ in $k$-nearest-neighbors

$k=1$

$k=7$

$k=15$

Bayes-optimal boundary given true generating model