**Similarity Estimation:**

- Applications for similarity or dissimilarity terms:
  - Retrieval Queries (Ranking, Range Queries)
  - Clustering
  - Model Training

- Similarity estimates reflect actual data similarity, i.e.:
  \[ s(x_1, x_2) > s(x_1, x_3) \Rightarrow x_1 \text{ more similar to } x_2 \text{ than to } x_3 \]

- Commonly used: Distance Measures
  - For \( x_1, x_2 \in \mathbb{R}^d \): \( L_p \)-norms
    \[ L_p(x_1, x_2) = \left( \sum_{i=1}^{d} |x_{1,i} - x_{2,i}|^p \right)^{1/p} \]

- Pitfalls:
  - Target similarity is ignored
  - Irrelevant features / Correlated features
  - Large influence of single dimensions
  - Arbitrarily large distances

**Our Solution:**

- Split the problem into similar (\( \text{sim} \)) and dissimilar (\( \text{dis} \)) object pairs
- Control the influence of a dimension \( i \) as probability in \([0,1]\) using a Bayes Estimate (BE):
  \[ \text{BE}(x_1, x_2) = \frac{\rho_{\text{sim}} \cdot P(x_1, x_2) \| \text{dis} \)}{\rho_{\text{dis}} \cdot P(x_1, x_2) \| \text{dis} \} + \rho_{\text{sim}} \cdot P(x_1, x_2) \| \text{sim} \]

- Global distance = Bayes Ensemble Distance (BED):
  \[ \text{BED}(x_1, x_2) = \frac{1}{d} \sum_{i=1}^{d} \text{BE}(x_1, x_2) \]

- More stable against outlier dimensions than a classical Naive Bayes Classifier:
  \[ \text{NB}(x_1, x_2) = \frac{1}{d} \prod_{i=1}^{d} \text{BE}(x_1, x_2) \]

**Difference Distributions:**

- Similar (\( \text{sim} \)) and dissimilar (\( \text{dis} \)) object pairs can form differentiable difference distributions
- Example distributions for 32-dimensional color
- Histogram Data of 34 class image dataset

**Feature Quality Assessment:**

- Ensemble method: Introduce weights
  - Relevance terms for variance difference of sim and dis
    \[ q_i = \sigma_{\text{sim}}^2 - \sigma_{\text{dis}}^2 = \text{avg}(x_{1,i}^2) - \text{avg}(x_{2,i}^2) \]
  - Include into ensemble:
    \[ \text{BED}(x_1, x_2) = \left( \sum_{i=1}^{d} q_i \right)^{-1} \sum_{i=1}^{d} q_i \cdot \text{BE}(x_1, x_2) \]

**Feature Space Improvement:**

- Dimensionality reduction
- Exploitation of correlated features

**Algorithm:**

**Input:** \( X \) with \( x_i \in \mathbb{R}^d \), SIM, DIS, target dimension \( d^* \)
- (1) Derive \( \Sigma_{\text{sim}} \) and \( \Sigma_{\text{dis}} \)
- (2) Compute feature transformation \( W \in \mathbb{R}^{d \times d^*} \)
- (3) Get weights \( q_i (i \in 1…d^*) \)

**Output:** Bayes Ensemble Distance
  \[ \text{BED}(x_1, x_2) = \left( \sum_{i=1}^{d} q_i \right)^{-1} \sum_{i=1}^{d} q_i \cdot \text{BE}(W^T x_1, W^T x_2) \]

**Conclusions on BEDs:**

- Balanced, adaptive distance measure
- Easily interpretable
- Applicable to various datasets (discrete class labels, pair-wise similarity labels, regression target functions)