Ranking on Data Manifolds

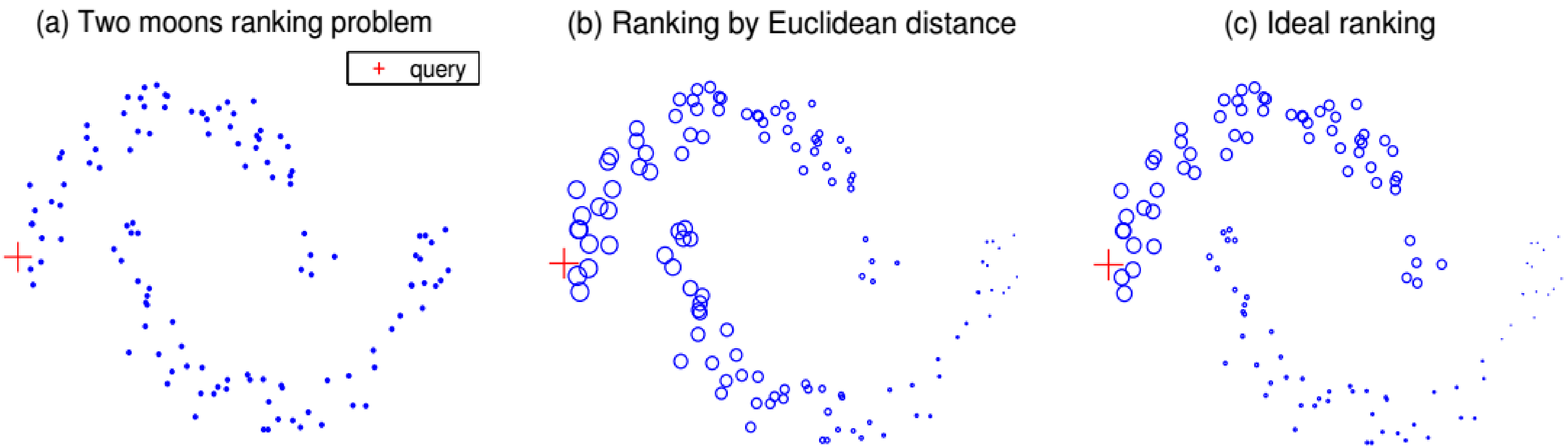

Figure 1: Ranking on the two moons pattern. The marker sizes are proportional to the ranking in the last two figures. (a) toy data set with a single query; (b) ranking by the Euclidean distances; (c) ideal ranking result we hope to obtain.

[Zhou *et al*. NIPS, 2003.]

$$\mathcal{X} = \{x_1, ..., x_q, x_{q+1}, ..., x_n\} \subset \mathbb{R}^m \qquad y = [y_1, .., y_n]^T$$

$f : \mathcal{X} \longrightarrow \mathbb{R}$: a ranking function which assigns to each point x_i a value f_i

$$f = [f_1, .., f_n]^T$$

1. Sort the pairwise distances among points in ascending order. Repeat connecting the two points with an edge according the order until a connected graph is obtained.
2. Form the affinity matrix W defined by $W_{ij} = \exp[-d^2(x_i, x_j)/2\sigma^2]$ if there is an edge linking x_i and x_j. Note that $W_{ii} = 0$ because there are no loops in the graph.
3. Symmetrically normalize W by $S = D^{-1/2} W D^{-1/2}$ in which D is the diagonal matrix with (i, i)-element equal to the sum of the i-th row of W.
4. Iterate $f(t+1) = \alpha S f(t) + (1 - \alpha) y$ until convergence, where α is a parameter in $[0, 1)$.
5. Let f_i^* denote the limit of the sequence $\{f_i(t)\}$. Rank each point x_i according its ranking scores f_i^* (largest ranked first).

$$f^* = (1 - \alpha)(I - \alpha S)^{-1} y$$

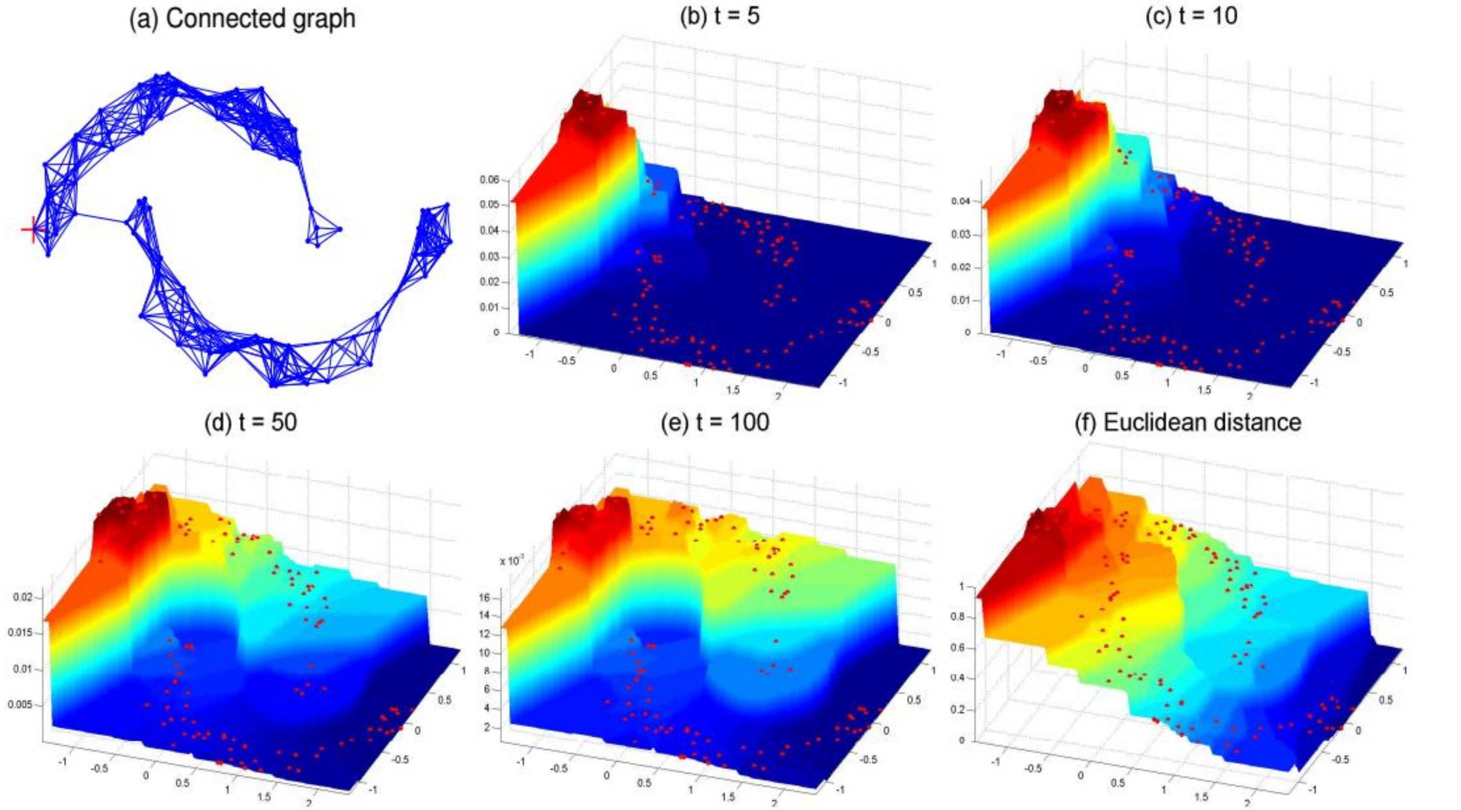

Figure 2: Ranking on the pattern of two moons. (a) connected graph; (b)-(e) ranking with the different time steps: $t = 5, 10, 50, 100$; (f) ranking by Euclidean distance.

Active Learning Methods

1. select the most relevant images

$$f^{+*}$$

2. select the most informative unlabeled images

$$\left|f_i^*\right| = \left|f_i^{+*} + f_i^{-*}\right|$$

Margin sample

3. selecting the inconsistent images which are also quite similar to the query

$$c(x_i) = f_i^{+*} - \left|f_i^{+*} + f_i^{-*}\right|$$

[He *et al*. Manifold-Ranking Based Image Retrieval. ACM MM, 2004.]

ERM

$$\arg\min_{f\in H_k} \frac{\pi}{n_P}\sum_{x_i\in X_P} l(f(x_i),+1) - l(f(x_i),-1) + \frac{1}{n_U}\sum_{x_j\in X_U} l(f(x_j),-1)$$

$$= \arg\min_{f\in H_k} -\frac{\pi}{n_P}\sum_{x_i\in X_P} f(x_i) + \frac{1}{n_U}\sum_{x_i\in X_U} l(f(x_i),-1)$$

$$l(f(x_i),+1) - l(f(x_i),-1) = -f(x_i)$$

$$= \arg\min_{f\in H_k} \widehat{R}_P + \widehat{R}_U$$

$$g(f,P,U) = -\frac{\pi}{n_P}\sum_{x_i\in X_P} f(x_i) + \frac{1}{n_U}\sum_{x_i\in X_U} l(f(x_i),-1) + \frac{\lambda}{2}\|f\|_{H_k}^2$$

$$\{Q^*, f^*\} = \arg\min_{Q,f} \widehat{R}_{P\cup Q_+} + \widehat{R}_{U\setminus Q_+}$$

$$\arg\min_{q^T 1_U = b,\, f\in H_k} \tilde{g}(f,P,U,q)$$

q: 被选中的概率

p: 被当前分类器预测为正类的概率

where $$\tilde{g}(f,P,U,q) = -\frac{\pi}{n_P}\sum_{x_i\in X_P} f(x_i) + \frac{1}{n_U}\sum_{x_j\in X_U} -q_j p_j \pi f(x_j) + (1-q_j p_j) l(f(x_j),-1) + \frac{\lambda}{2}\|f\|_{H_k}^2$$

$$\tilde{g}(f,P,U,q) = -\frac{\pi}{p}\sum_{x_i\in X_P} f(x_i) + \frac{1}{u}\sum_{x_j\in X_U} -q_j p_j \pi f(x_i) + (1-q_j p_j) l(-f(x_j)) + \frac{\lambda}{2}\|f\|_{H_k}^2$$

$$= \left\{-\frac{\pi}{p}\sum_{x_i\in X_P} f(x_i) + \frac{1}{u}\sum_{x_j\in X_U} l(-f(x_j)) + \frac{\lambda}{2}\|f\|_{H_k}^2\right\} + \frac{1}{u}\sum_{x_j\in X_U} -q_j p_j \pi f(x_j) - q_j p_j l(-f(x_j))$$

$$= g(f,P,U) - \frac{1}{u}\sum_{x_j\in X_U} q_j p_j \left(\pi f(x_j) + l(-f(x_j))\right)$$

1. Fixed f, optimize q

$$\underset{q^T 1=b,\, f\in H_k}{\arg\min}\ \tilde{g}\left(f,P,U,q\right)$$

$$= \underset{q^T 1=b}{\arg\min}\ \tilde{g}\left(f^*,P,U,q\right)$$

$$= \underset{q^T 1=b}{\arg\min}\ g(f^*,P,U) - \frac{1}{u}\sum_{x_j\in X_U} q_j p_j \left(\pi f^*(x_j) + l(-f^*(x_j))\right)$$

$$= \underset{q^T 1=b}{\arg\min}\ g(f^*,P,U) - \frac{1}{u}\sum_{x_j\in X_U} q_j p_j \pi f^*(x_i) - \frac{1}{u}\sum_{x_j\in X_U} q_j p_j l(-f^*(x_j))$$

$$= \underset{q^T 1=b}{\arg\min}\ g(f^*,P,U) + \langle q\circ p\rangle^T 1_U \left(\pi \widehat{R}_{U,P} - (1-\pi)\widehat{R}_{U,N}\right)$$

2. Fixed q, optimize f

Problem:

1. 如果 batch 中没有正例被标注出来，PU集合都不会发生改变，会陷入死循环。
2. 可解释性

Vectorization

$$f(x_j) = \sum_{i:x_i \in D} \alpha_i \phi(x_i) = \alpha^T \phi(x_i)$$

$$\max_{q^T 1_U = b} \sum_{x_j \in X_U} q_j p_j \left\{ \pi\left(\alpha^T \phi(x_j)\right) + \frac{1}{4}\left(\left(\alpha^T \phi(x_j)\right)^2 + 2\left(\alpha^T \phi(x_j)\right) + 1\right) \right\}$$

$$= (\pi + \frac{1}{2}) \langle q \circ p \rangle^T K_{U,D} \alpha + \frac{1}{4} \alpha^T K_{U,D}^T \langle q \circ p \rangle \langle q \circ p \rangle^T K_{U,D} \alpha + \frac{1}{4} \langle q \circ p \rangle^T 1_U$$