Q2) Semi Supervised Setting

Objective : How can we learn a graph matrix integrating the label information.

→ There are many techniques that are in the literature to learn a ~~l~~ graph with semi-supervised setting. But most of them follow the same step.

(1) Learn the graph with any of the framework which exist by not including the label information

(2) Now using the graph so obtained, find the missing labels. There are different techniques out there to assign a label, such as label propagation, Random Walk etc

→ But the problem with all these methods is that these they are not taking these available label informations to learn a graph which can be very informative

→ There was a paper published by 'Lianshong Zhuang' {Label Information Guided Graph construction for Semi Supervised Learning} where he motivated the

2

need for integrating label information. The key idea he proposed was by enforcing the weights of edges between labelled samples of different class to be zero

→ We can include this idea as an additional constraint while learning the graph using the method given in my solution to question 1

Previous Objective func:

$$\min_{W} \left[\|W \circ Z\|_{1,1} - \alpha \mathbf{1}^T \log(W \mathbf{1}) + \beta \|W\|_F^2 \right]$$

Modification:

→ We are been given with $\{x_i, y_i\}_{i=1}^{k}$ & $\{x_j\}_{j=k+1}^{n}$ & $y_i \in \{0, 1\}$

→ We can impose a constraint on labelled data point such that

$$W_{ij} = 0 \quad \forall (i,j) \ \text{s.t} \ y_i \neq y_j$$

→ This will ensure that 2 ~~classes~~ nodes which are not having the same class will have no connection between them.

Algorithm :

1) Change $y_i \in \{0, 1\}^k$ to $y_i \in \{-1, 1\}^k$

2) Assign unlabeled y_i to zero

3) New optimization step :

$$\min \left[\|W \circ Z\|_{1,1} - \alpha \, 1^T \log(W 1) + \beta \|W\|_F^2 \right]$$

s.t $W_{ij} = 0$ where $|y_i - y_j| = 2$

This condition will ensure that unlabelled data (now assigned a label 0) will be independent of the new constraint & for the ~~ls~~ labels which are different, the respective edge will be forced to zero.

The final objective can be written as

$$\min \left[\|W \circ Z\|_{1,1} - \alpha \, 1^T \log(W 1) + \beta \|W\|_F^2 + \sum_{(i,j)} \gamma \left(2 - |y_i - y_j|\right) W_{ij} \right]$$

4) Now as we have obtained the graph matrix W, we can make use of that and learn the unlabelled data

We can make use of ~~label~~ the classic label propagation for this.

Label Propagation ~~(by z.)~~

→ ~~Split~~ Take labelled Y as $Y_L = \{y_1 \cdots y_L\}$

→ Unlabelled $Y = Y_U = \{y_{l+1}, y_{l+2} \cdots y_n\}$

→ Construct a probabilistic transition matrix T s.t

$$T_{ij} = P(j \to i) = \frac{w_{ij}}{\sum_{k=1}^{n} w_{kj}}$$

where $P(j \to i)$ is the prob. of propagating label of node j to node 'i'

So the algorithm can be written as

Step 1) Assign new Y, $Y \leftarrow TY$

Step 2) Normalize Y to maintain class probability

Step 3) Clamp labelled data & repeat step 1 and 2 until convergence

→ Because of this clamping, we can see only unlabelled Y_U will be changing.

→ We can ~~ya~~ write matrix $\bar{T} = \begin{bmatrix} \bar{T}_{ll} & \bar{T}_{lu} \\ \bar{T}_{ul} & \bar{T}_{uu} \end{bmatrix}$ form

⇒ Update of $Y_U \leftarrow \bar{T}_{uu} Y_u + \bar{T}_{ul} Y_L$

$\Rightarrow$ It can be shown that

$$Y_u \leftarrow (1 - \bar{T}_{uu})^{-1} \bar{T}_{ul} Y_L$$

which will be a fixed point & thus it converges.