October 10, 2018

End-to-en differentiable Label Propagation library, implemented using TensorFlow.

The underlying idea is that it is possible to spread labels from one (or a few nodes) in an undirected graph to all $n$ nodes, by minimising the following cost function, defined over a node labeling $f \in R^n$:

$$E(f) = \sum_{i \in L} (f[i] - y[i])^2 + \mu \sum_i \sum_j W_i j (f[i] - f[j])^2 + \mu\epsilon \sum_i f_i^2. \qquad (1)$$

- The term $\sum_{i \in L} (f[i] - y[i])^2$ enforces consistency of labeled nodes (i.e. those in $L$) with a gold labeling $y$.

- The term $\sum_i \sum_j W_{ij} (f[i] - f[j])^2$ enforces that, given two nodes that are connected in the undirected graph (with weight $W_{ij} = W_{ji} > 0$), they are associated to a similar labeling.

- The term $\sum_i f[i]^2$ is a $L2$ regulariser.

  Since the cost function $E(f)$ is quadratic, it has one closed-form solution for $\mu > 0$ and $\epsilon > 0$. Furthermore, it is possible to backpropagate the error resulting from the propagation process, back to the graph structure encoded by the adjacency graph $W$.