Learning User Perceived Clusters with Feature-Level Supervision

Authored by:

Ting-Yu Cheng
Guiguan Lin
xinyang gong
Kang-Jun Liu
Shan-Hung Wu

Abstract

Semi-supervised clustering algorithms have been proposed to identify data clusters that align with user perceived ones via the aid of side information such as seeds or pairwise constrains. However, traditional side information is mostly at the instance level and subject to the sampling bias, where non-randomly sampled instances in the supervision can mislead the algorithms to wrong clusters. In this paper, we propose learning from the feature-level supervision. We show that this kind of supervision can be easily obtained in the form of perception vectors in many applications. Then we present novel algorithms, called Perception Embedded (PE) clustering, that exploit the perception vectors as well as traditional side information to find clusters perceived by the user. Extensive experiments are conducted on real datasets and the results demonstrate the effectiveness of PE empirically.

1 Paper Body

x2
x1
100
50 40
OKM OKM* OKM*+LPE OKM*+NPE
80
30
60
20
40
10

20
0 0
0.2
0.4
0.6
F?Score
0.8
1
0 0
OKM* OKM*+LPE OKM*+NPE
0.2
0.4
0.6
F?Score
0.8
1

2 References

NA