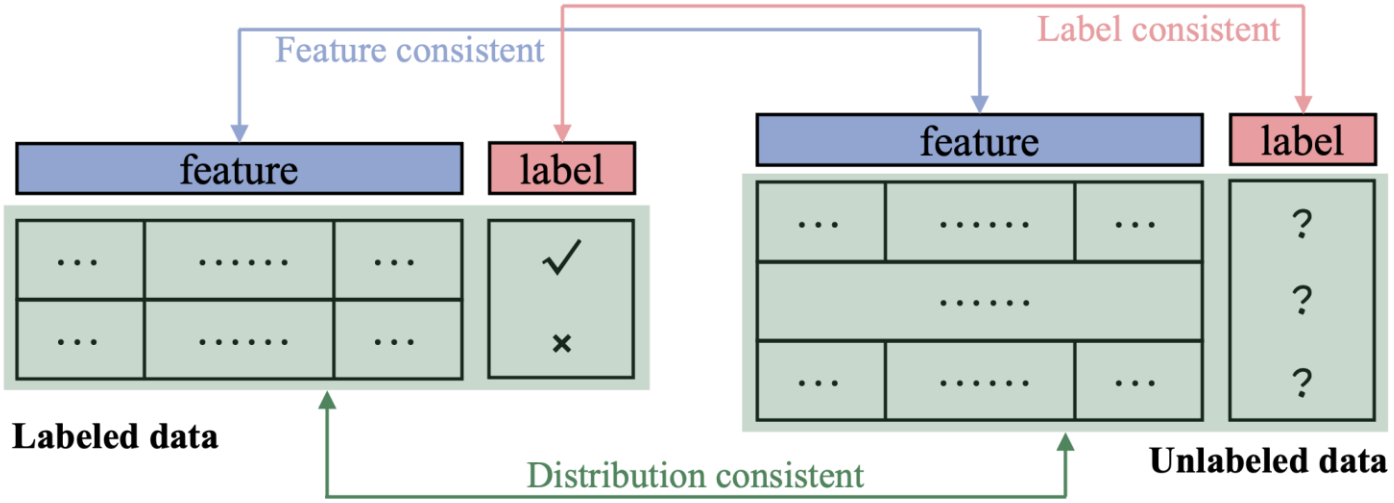
Robust Semi-Supervised Learning in Open Environments

Lan-Zhe GUO, Lin-Han JIA, Jie-Jing SHAO, Yu-Feng LI

Frontiers of Computer Science, DOI: 10.1007/s11704-024-40646-w

Background

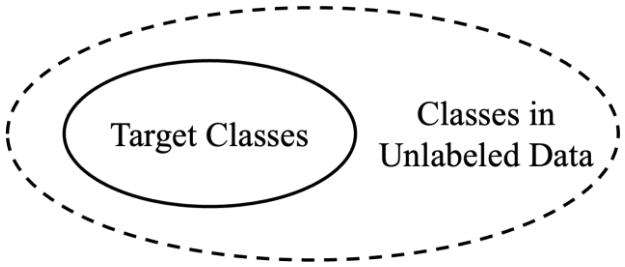
- Semi-Supervised learning typically assume close environments where important factors (e.g., label, feature, distribution are consistent between labeled and unlabeled data.



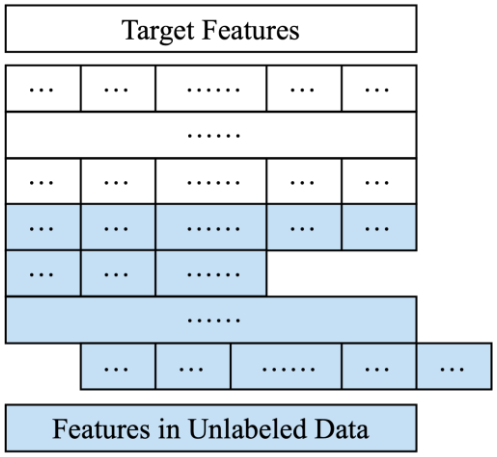
- However, many real-world tasks involve open environments, where class label set, feature space, and data distribution could be inconsistent. This paper give a brief review of advances in this line of research.

Main Contributions

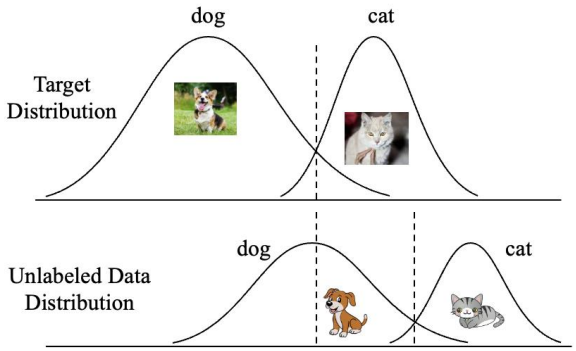
- Contributions:
  - This paper briefly introduces some advances of SSL in open environments, focusing on techniques concerning label space mismatch, feature space mismatch and distribution mismatch between labeled and unlabeled data.
  - A new benchmark is provided to help evaluate the robustness of SSL in open environments. Some novel performance measures and an open-sourced SSL toolkit LAMDA-SSL are provided.



Label space mismatch



Feature space mismatch



Distribution mismatch