Learning with Weak Supervision

（弱监督机器学习范式）

Min-Ling Zhang (张敏灵)

PALM Group, School of Computer Science and Engineering,

MOE Key Laboratory of Computer Network & Information Integration, Southeast University, China

July 23, Hangzhou

Big Data

Essential Goal

Turn data into information and knowledge, so as to support sound decision making

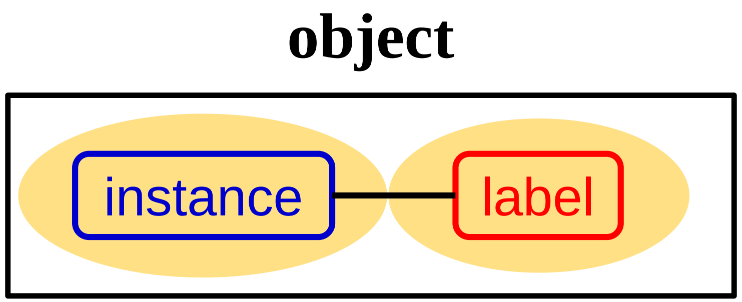
Key Techniques

Cloud Computing → Managing Data

Crowdsourcing → Collecting Data

Machine Learning → Analyzing Data

Traditional Supervised Learning

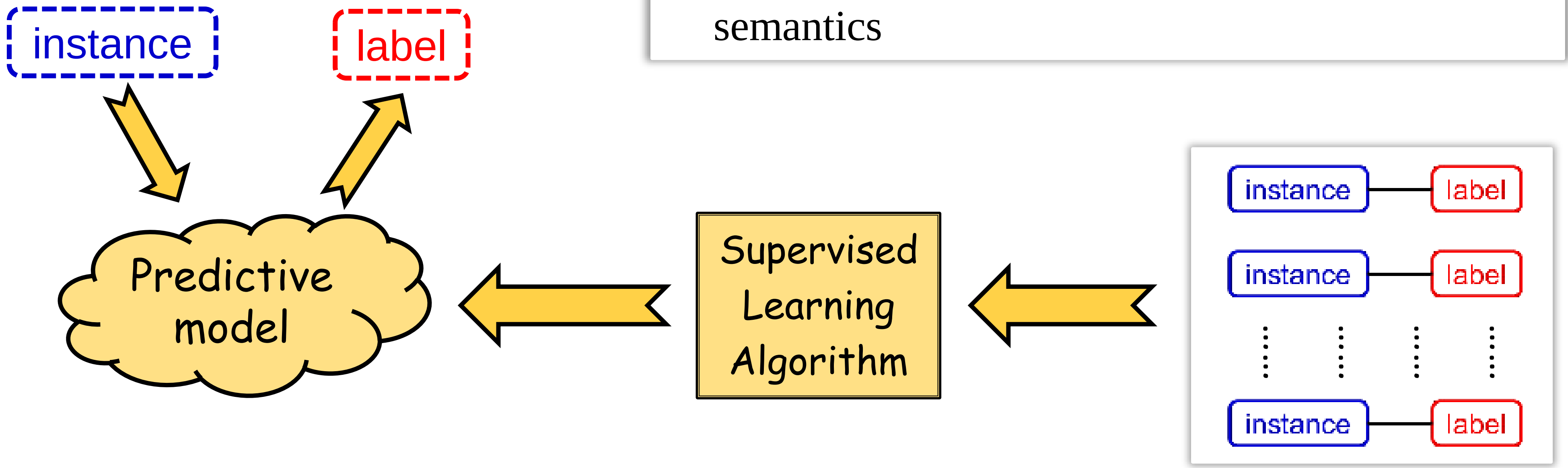


Input Space

represented by a single instance (feature vector) characterizing its properties

Output Space

associated with a single label characterizing its semantics



Basic Assumption: Strong Supervision

label ⟺ supervision information

Key factor for successful learning

(encoding *semantics* and *regularities* for the learning problem)

Strong supervision assumption

- Sufficient labeling
  abundant labeled training data are available
- Explicit labeling
  object labeling is unique and unambiguous

But, Supervision Is Usually Weak

Difficult to have!

Constrained by:

- Limited resources
- Physical environment
- Problem properties
- ……

Strong supervision (sufficient & explicit) → **Strong generalization ability**

In practice, we usually have to learn with weak supervision

Learning with Weak Supervision

- ✓ **Insufficient labeling**

  Labeled Data + Unlabeled Data

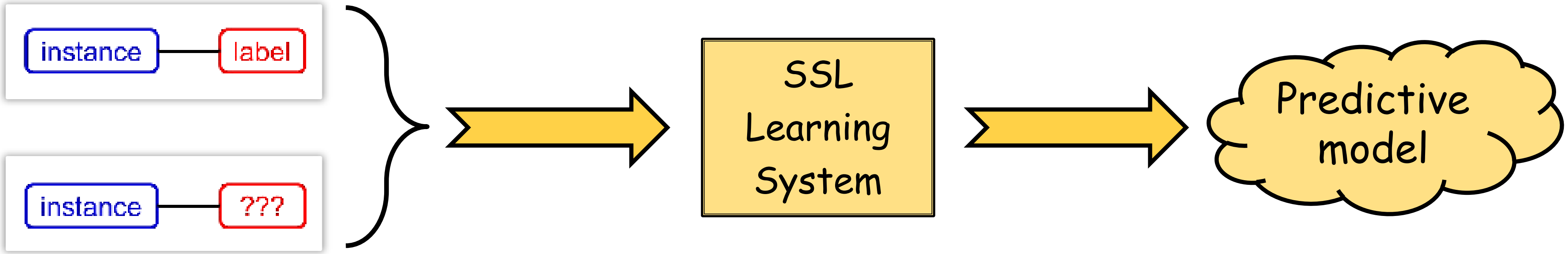
- ✓ **Non-Unique labeling**

  Multi-Label Data (labeling with multiple valid labels)

- ✓ **Ambiguous labeling**

  Partial-Label Data (labeling with multiple candidate labels)

Semi-Supervised Learning (SSL)



Major paradigm in exploiting unlabeled data to improve generalization performance, without human interventions

- Generative methods [Miller & Uyar, NIPS'97] [Nigam et al., MLJ00]
- S3VMs [Joachims, ICML'99] [Chapelle & Zien, AIStats'05] [Grandvalet & Bengio, NIPS'05]
- Graph-based methods [Zhu et al., ICML'03] [Zhou et al, NIPS'04] [Belkin et al., JMLR06]
- Disagreement-based methods [Blum & Mitchell, COLT'98] [Zhou & Li, KAIS10]

Multi-Label Objects

Russia to Ensure High Level of Security at 2018 FIFA World Cup

SPORT 21:53 17.04.2017

Topic: Russia hosts FIFA World Cup 2018

1 286 3 5

Russia will provide a proper level of security on the streets and in public places during the 2018 FIFA World Cup in the country, Vladimir Markin, the head of the Russian Football Union's security committee, said Monday.

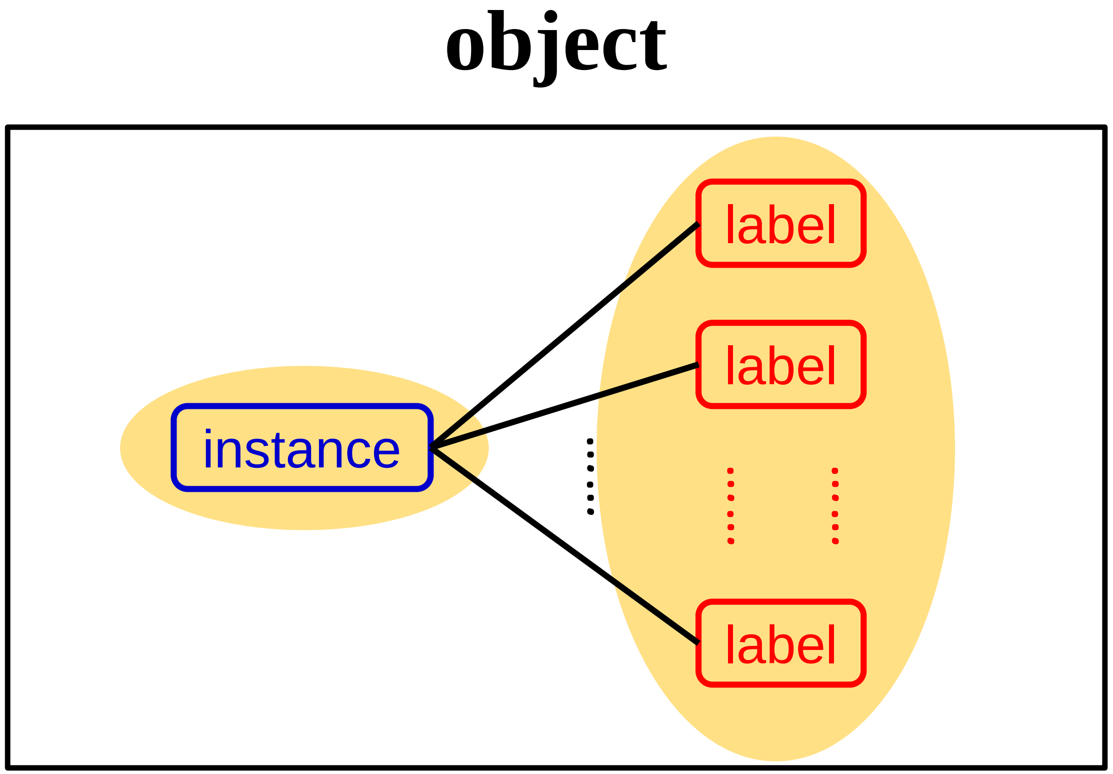
MOSCOW (Sputnik) — Earlier in the day, Russian President Vladimir Putin, amid preparations for the international football event, signed a federal law that will toughen administrative responsibility for football fans and introduce penalties for gross violations of rules of conduct during sporting events.

> *"The security on the streets and in other public places of the cities where the matches will be held is going to be provided on a proper level. Not only the law enforcement, but also private security services, volunteers and responsible citizens will make every effort for it," Markin told R-Sport.*

- Sports
- Europe
- Economics
- Travel
- Government
- ……

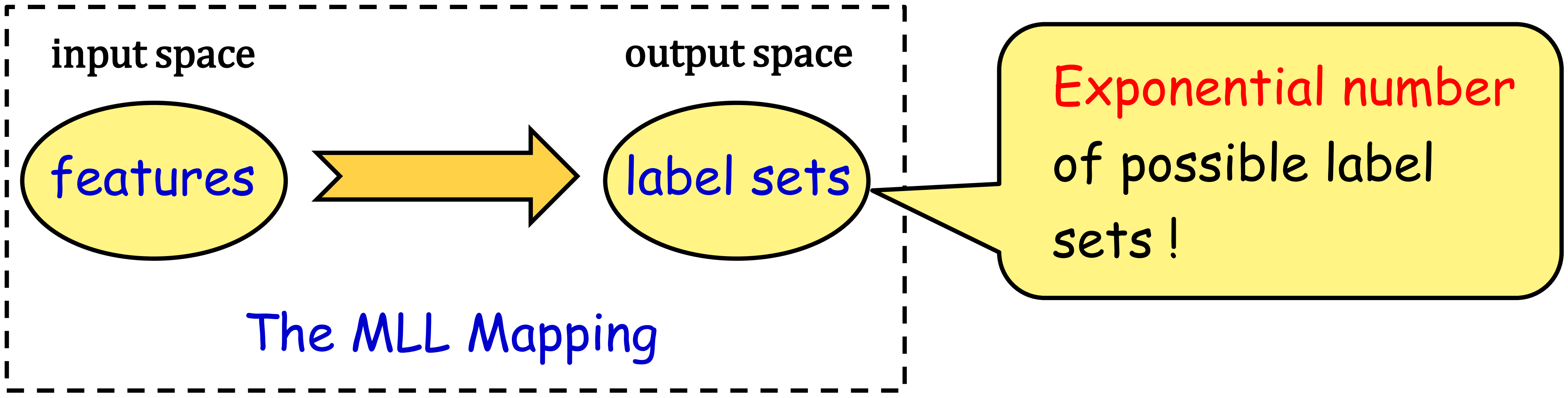
Multiple labels

Multi-Label Learning (MLL)



Multi-Label Learning (MLL)

Major Challenge of MLL



q=5 ➔ 32 label sets

q=10 ➔ ~1k label sets

q=20 ➔ ~1M label sets

……

Supervision Info.

- ☐ Individually strong
- ☐ But, globally weak !

Partial Label

Appreciator A ---------->

Appreciator B ---------->

Appreciator C ---------->

----------> Picasso style ×

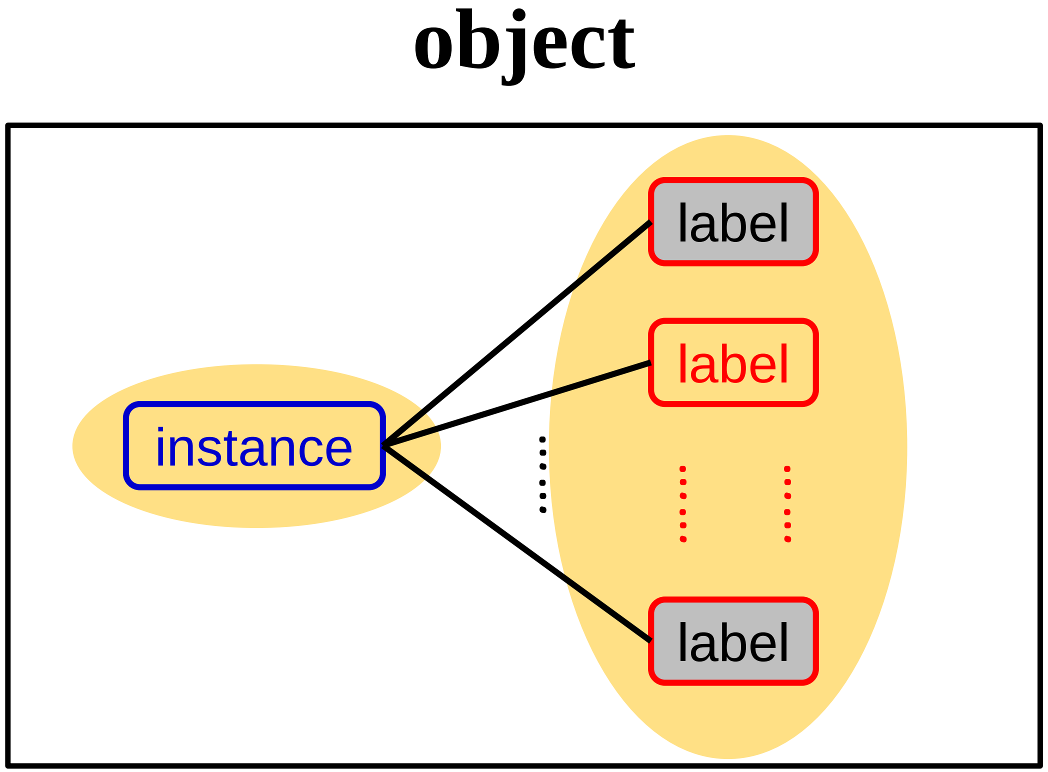
----------> Monet style ×

----------> van Gogh style √

Widely exist in real-world applications

- Computer vision [Cour et al., JMLR11] [Tang & Zhang, AAAI'17]
- Image classification [Zeng et al., CVPR'13] [Chen et al., CVPR'13]
- Learning from crowds [Raykar et al., JMLR10] [Yu & Zhang, MLJ17]
- Ecoinformatics [Liu & Dietterich, NIPS'12] [Zhang & Yu, IJCAI'15]
- ……

Partial-Label Learning (PLL)



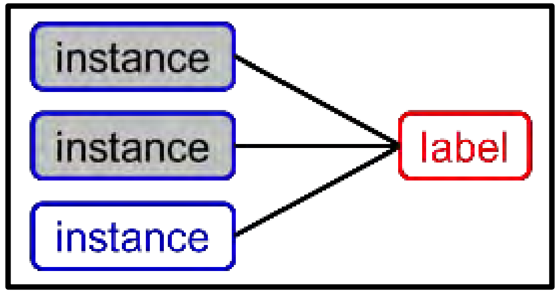
- Each object is associated with multiple candidate labels
- Only one of the candidate label is the unknown ground-truth label

Partial-Label Learning (PLL)

Other Scenarios Widely Exist

multi-instance learning

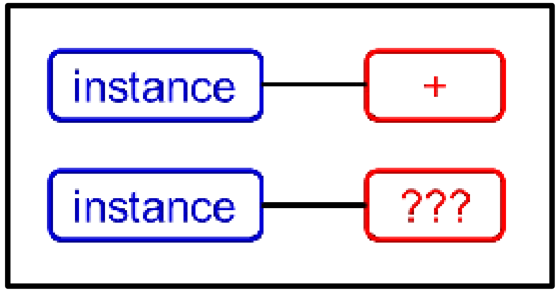
[Dietterich et al., AIJ97] [Foulds & Frank, KER10] [Amores, AIJ13]



ambiguous labeling

PU learning

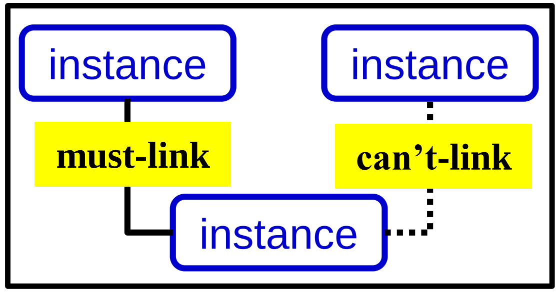
[Liu et al., ICML'02] [Liu et al., ICDM'03] [Li et al., ACL'10]



insufficient labeling

learning with constraints

[Wagstaff et al., ICML'01] [Basu et al., CRCBook08]



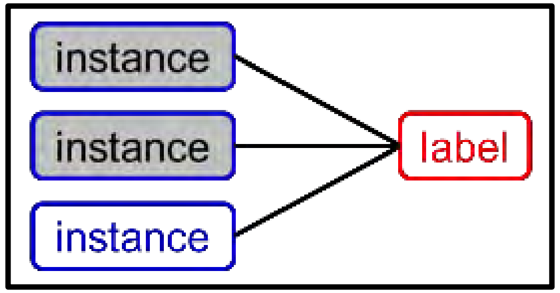
non-unique labeling

…………

Other Scenarios Widely Exist

multi-instance learning

[Dietterich et al., AIJ97] [Foulds & Frank, KER10] [Amores, AIJ13]

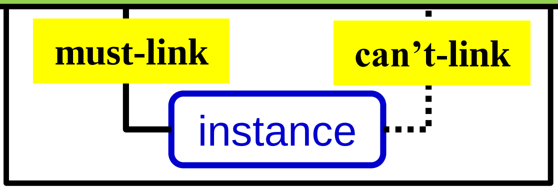


ambiguous labeling

Learning with Weak Supervision

Framework + Model + Utilization

le



[Wagstaff et al., ICML'01] [Basu et al., CRCBook08]

…………

Thanks!