↳ Supervised Learning (all examples labelled) | ↳ Unsupervised Learning (No examples labelled) | ↳ Reinforcement Learning (deferred labels)

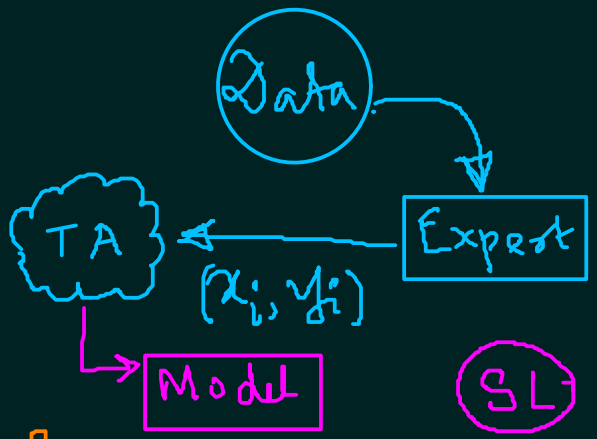
→ Real World ⇒ few labelled examples ⊕ many unlabelled data points } → Can these help?

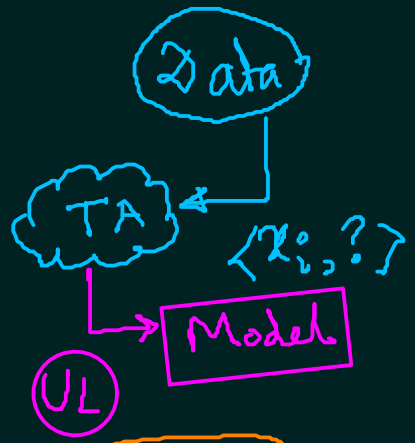
Examples:

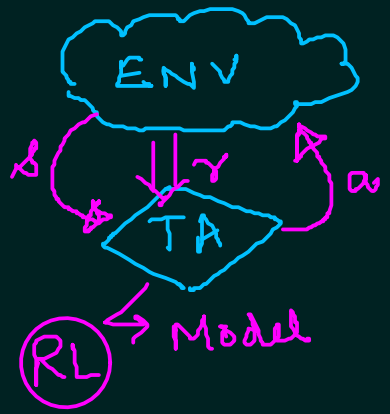
↳ Web-doc classify
↳ Computer Vision → Medical Imaging
↳ Computational Biology

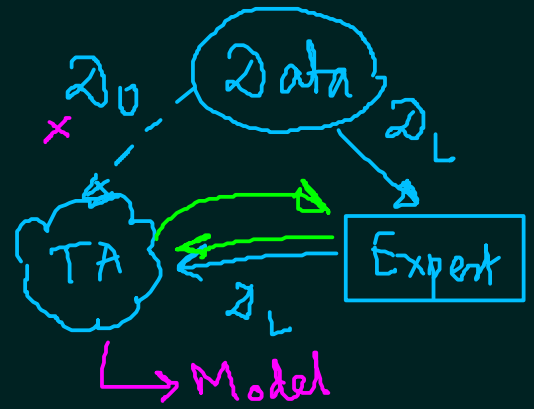
Semi-Supervised Learning | Active (Interactive) Learning

(distribution of data helps!)

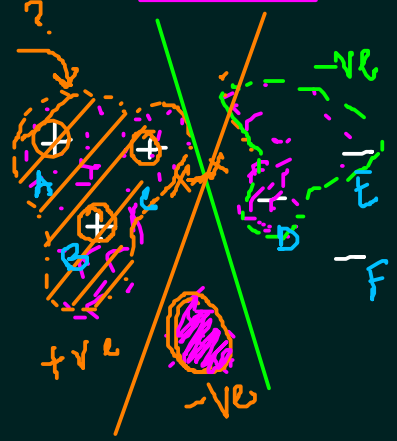


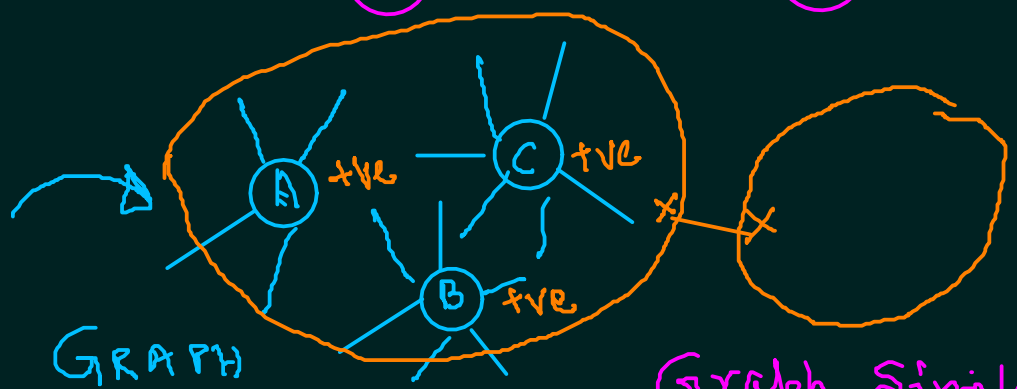






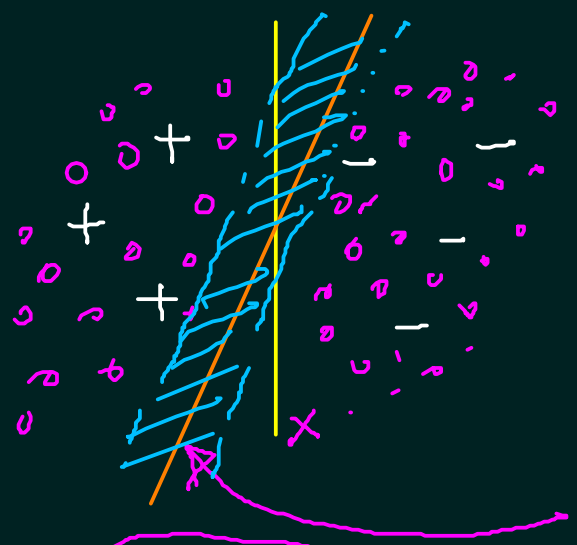
$[\mathcal{D} = \mathcal{D}_L \cup \mathcal{D}_U \quad \langle x, y \rangle \quad \langle x, ? \rangle]$





CCTV monitoring (webcams)

Graph Similarity based Method for SSL



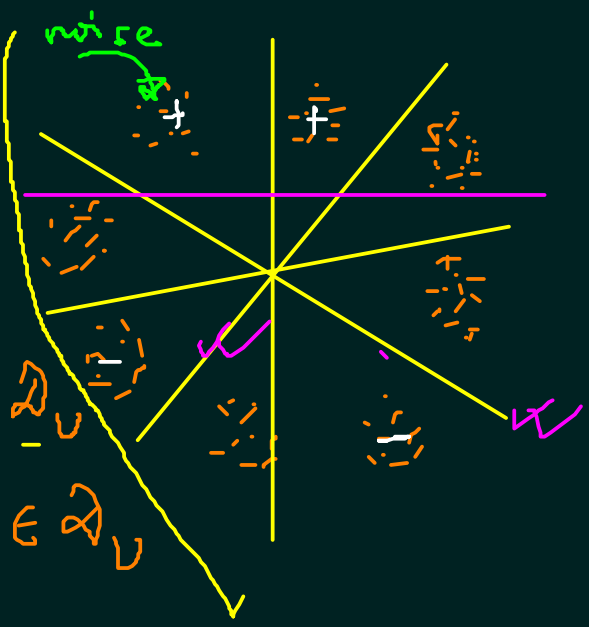
$$\text{argmin}_{w} \ \frac{1}{2} w w^T$$

$$\text{s.t.} \quad * \ y_i w \cdot x_i \geq 1 \quad \forall i \in \mathcal{D}_L$$

$$\hat{y}_u \cdot w \cdot x_u \geq 1 \quad , \quad \forall u \in \mathcal{D}_U$$

$$\hat{y}_u \in \{+1, -1\} \quad , \quad \forall u \in \mathcal{D}_U$$

iid
μ, σ



Transductive SVM (Joachim '99)

(NP-hard)

Heuristic / Branch-&-Bound

Assume: large margin separability

Co-training (Blum Mitchell '98)

$$\text{Argmin}_{w} \ \frac{1}{2} w w^T + C \sum_l \xi_l + C \sum_u \hat{\xi}_u$$

$$\text{s.t.} \quad y_l w x_l \geq 1 - \xi_l \quad , \forall l \in \mathcal{D}_L$$

$$(i) \quad \hat{y}_u w x_u \geq 1 - \hat{\xi}_u \quad , \forall u \in \mathcal{D}_U$$

$$(ii) \quad \hat{y}_u \in \{+1, +1\} \quad , \forall u \in \mathcal{D}_U$$

Webpage → Faculty / Student

$x_1 \leftarrow$ links inside webpage [My advisor] ← Students / collaborator

$x_2 \leftarrow$ text description ← I am teaching CS6015

$\langle x_1, x_2 \rangle$ → $c_1(x_1) = c_2(x_2) = c^*(x)$

c_1 c_2

c_1 → P_1 → str
P_2 → fac

c_2 → P_2 fac
P_4 fac

Itr

P_1 P_2 P_3 P_4

→ Web-doc classify

SSL → Graph Similarity Based
Transductive SVM
Co training

Co training (Consensus)

SVM (slack)
Tr. SVM
Act. SVM

▷ Active Learning :

N points (unlabelled) → $O(\log_2 N)$

$O\left(\frac{1}{\epsilon}\right)$ ↑ ϵ (Sanjay Dasgupta 09)

Binary search

Exp. improve

→ Density Based Sampling
→ Uncertainty Based Sampling
→ Ensembling ways Sampling

(Sanj 2008)

Better than R. Sampl

$W_1 D + U W_2 \geq$ ve

ACTIVE SVM