Exploiting cluster-structure to predict the labeling of a graph

New Challenges in Theoretical Machine Learning: Learning with Data-dependent Concept Spaces

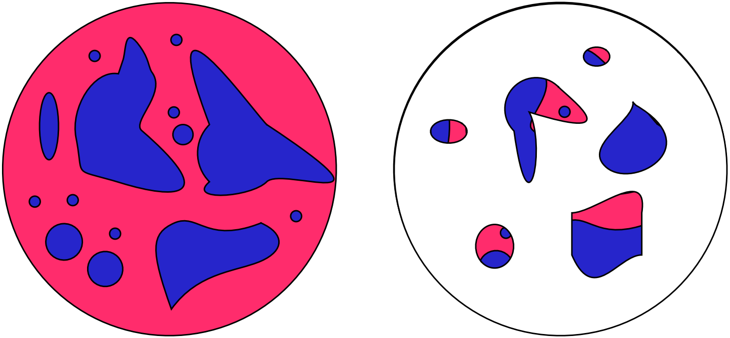
Mark Herbster

University College London
Department of Computer Science

13 December, 2008

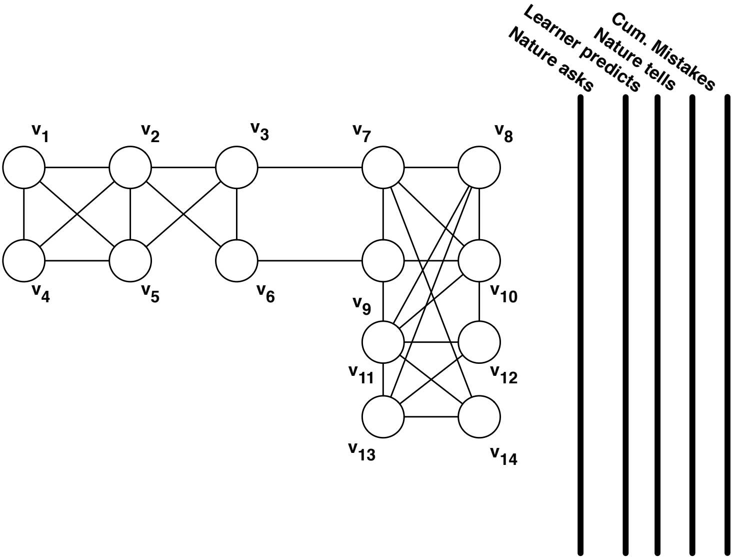
Overview

1. Give perceptron-like algorithm for graph label prediction
2. Improve on Perceptron bound when cluster-structure

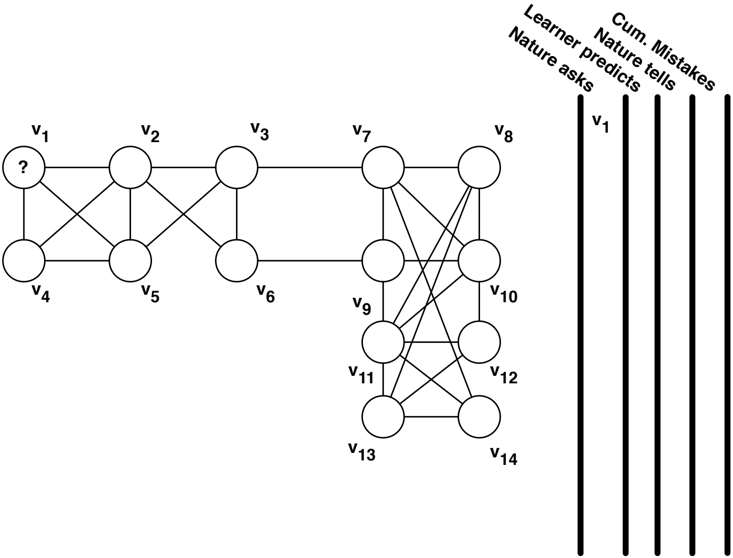


“Default” Assumption – *“Cluster” Assumption*

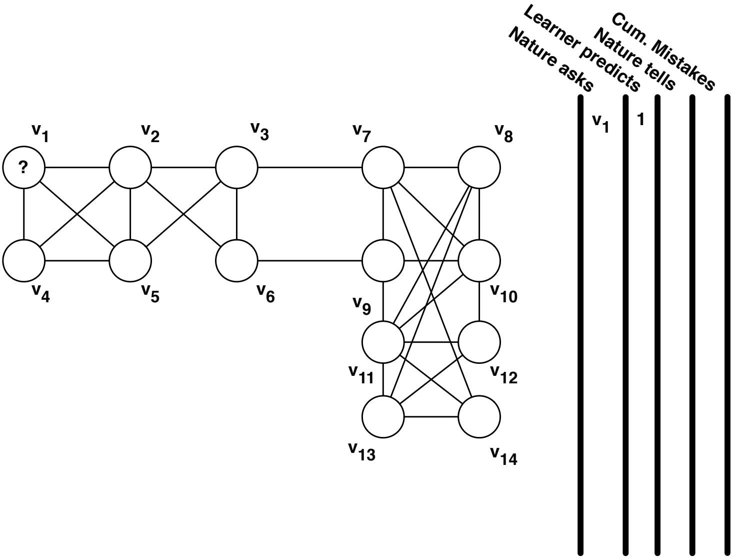
A prediction game



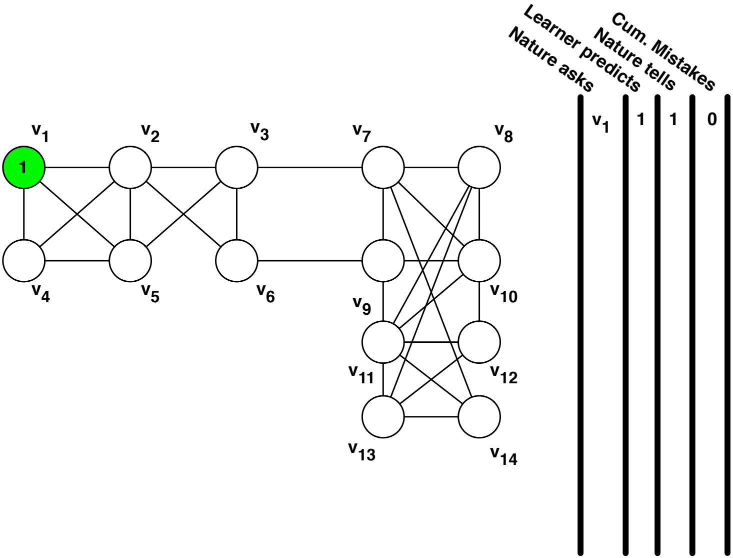
A prediction game



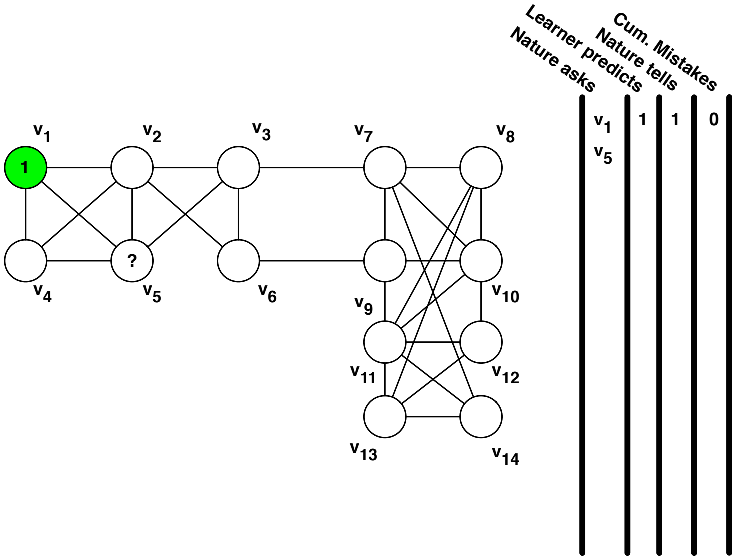
A prediction game



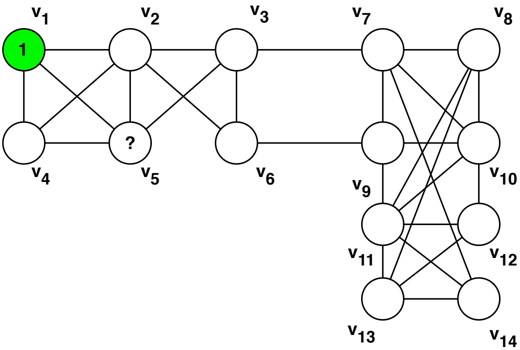
A prediction game



A prediction game

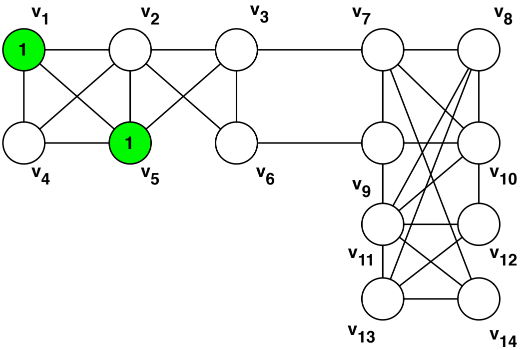


A prediction game



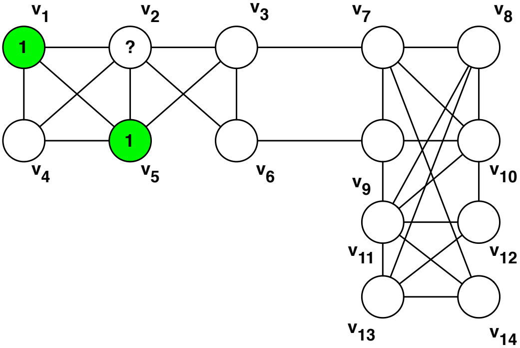
Nature asks	Learner predicts	Nature tells	Cum. Mistakes
v_1	1	1	0
v_5	1		

A prediction game



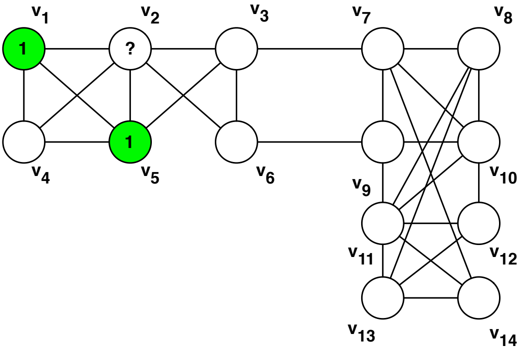
Nature asks	Learner predicts	Nature tells	Cum. Mistakes
v_1	1	1	0
v_5	1	1	0

A prediction game



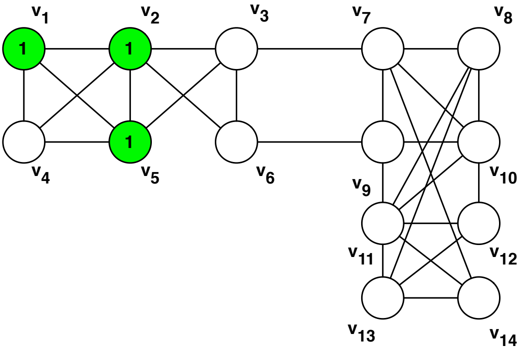
Nature asks	Learner predicts	Nature tells	Cum. Mistakes
v_1	1	1	0
v_5	1	1	0
v_2			

A prediction game



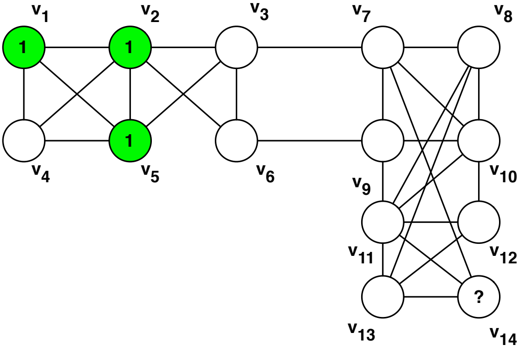
Nature asks	Learner predicts	Nature tells	Cum. Mistakes
v_1	1	1	0
v_5	1	1	0
v_2	1		

A prediction game



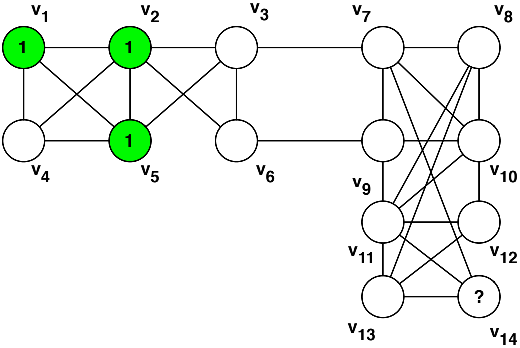
Nature asks	Learner predicts	Nature tells	Cum. Mistakes
v_1	1	1	0
v_5	1	1	0
v_2	1	1	0

A prediction game



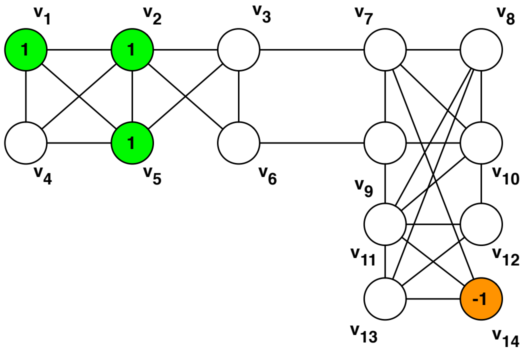
Nature asks	Learner predicts	Nature tells	Cum. Mistakes
v_1	1	1	0
v_5	1	1	0
v_2	1	1	0
v_{14}			

A prediction game



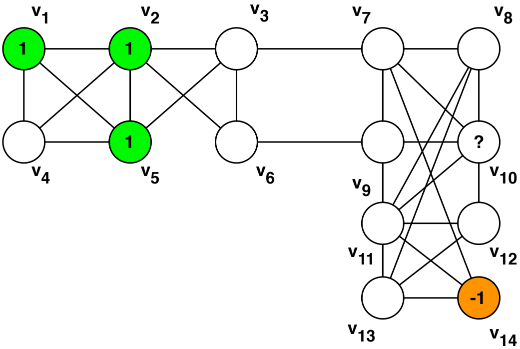
Nature asks	Learner predicts	Nature tells	Cum. Mistakes
v_1	1	1	0
v_5	1	1	0
v_2	1	1	0
v_{14}	1		

A prediction game



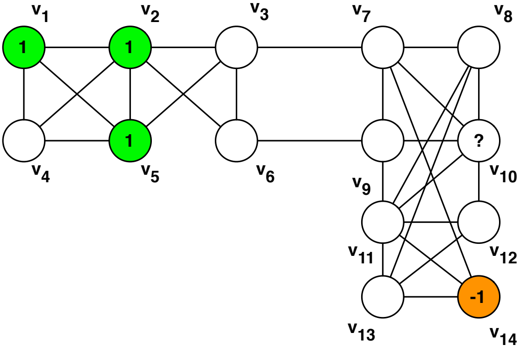
Nature asks	Learner predicts	Nature tells	Cum. Mistakes
v_1	1	1	0
v_5	1	1	0
v_2	1	1	0
v_{14}	1	-1	1

A prediction game



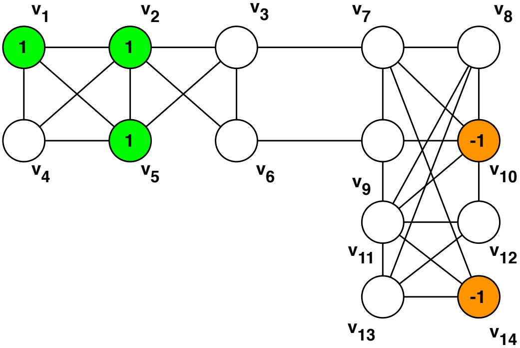
Nature asks	Learner predicts	Nature tells	Cum. Mistakes
v_1	1	1	0
v_5	1	1	0
v_2	1	1	0
v_{14}	1	-1	1
v_{10}			

A prediction game



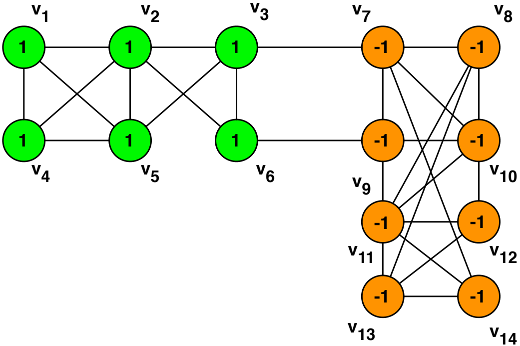
Nature asks	Learner predicts	Nature tells	Cum. Mistakes
v_1	1	1	0
v_5	1	1	0
v_2	1	1	0
v_{14}	1	-1	1
v_{10}	-1		

A prediction game



Nature asks	Learner predicts	Nature tells	Cum. Mistakes
v_1	1	1	0
v_5	1	1	0
v_2	1	1	0
v_{14}	1	-1	1
v_{10}	-1	-1	1

A prediction game



Nature asks	Learner predicts	Nature tells	Cum. Mistakes
v_1	1	1	0
v_5	1	1	0
v_2	1	1	0
v_{14}	1	-1	1
v_{10}	-1	-1	1
v_6	1	1	1
v_{11}	-1	-1	1
v_{13}	-1	-1	1
v_9	-1	-1	1
v_3	1	1	1
v_7	-1	-1	1
v_{12}	-1	-1	1
v_8	-1	-1	1
v_4	1	1	1

Perceptron Bound (Novikoff)

Theorem [Novikoff]:

Given a sequence $\{(\boldsymbol{x}_t, y_t)\}_{t=1}^{\ell} \subseteq \mathcal{H} \times \{-1, 1\}$ then the mistakes of the perceptron are bounded by

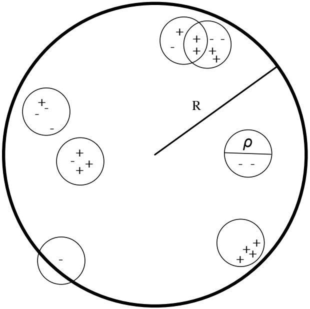
$$M \leq \|\boldsymbol{u}\|^2 R$$

with $R = \max_t(\|\boldsymbol{x}_t\|^2)$ for all $\boldsymbol{u} \in \mathcal{H}$ such that

$$\langle \boldsymbol{u}, \boldsymbol{x}_t \rangle y_t \geq 1$$

for $t = 1, \ldots, \ell$.

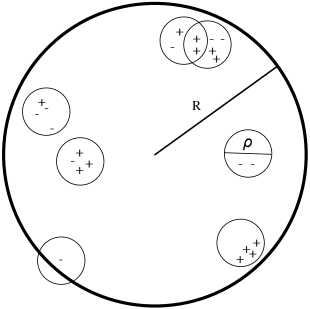
Pounce Bound Motivation



Input space X of radius R with cover number $\mathcal{N}(X, \rho) = 7$.

- Bounds to be dependent on structure of input space X.
- Novikoff is only dependent on X through radius R.
- Expectation is that a typical ambient input space is only sparsely populated (cf manifold/cluster hypotheses).
- Pounce will depend on the cover of X.
- In particular the number of balls $\mathcal{N}(X, \rho)$ of diameter ρ.

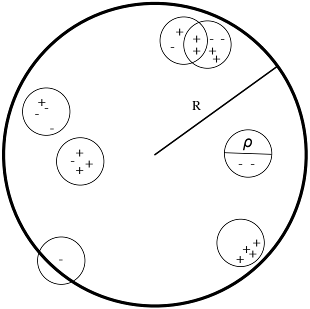
Pounce Bound Motivation



Input space X of radius R with cover number $\mathcal{N}(X, \rho) = 7$.

- Bounds to be dependent on structure of input space X.
- Novikoff is only dependent on X through radius R.
- Expectation is that a typical ambient input space is only sparsely populated (cf manifold/cluster hypotheses).
- Pounce will depend on the cover of X.
- In particular the number of balls $\mathcal{N}(X, \rho)$ of diameter ρ.

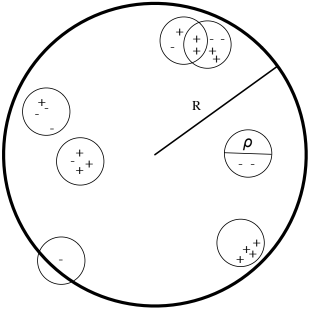
Pounce Bound Motivation



Input space X of radius R with cover number $\mathcal{N}(X, \rho) = 7$.

- Bounds to be dependent on structure of input space X.
- Novikoff is only dependent on X through radius R.
- Expectation is that a typical ambient input space is only sparsely populated (cf manifold/cluster hypotheses).
- Pounce will depend on the cover of X.
- In particular the number of balls $\mathcal{N}(X, \rho)$ of diameter ρ.

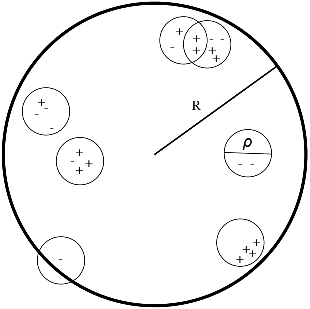
Pounce Bound Motivation



Input space X of radius R with cover number $\mathcal{N}(X, \rho) = 7$.

- Bounds to be dependent on structure of input space X.
- Novikoff is only dependent on X through radius R.
- Expectation is that a typical ambient input space is only sparsely populated (cf manifold/cluster hypotheses).
- Pounce will depend on the cover of X.
- In particular the number of balls $\mathcal{N}(X, \rho)$ of diameter ρ.

Pounce Bound Motivation



Input space X of radius R with cover number $\mathcal{N}(X, \rho) = 7$.

- Bounds to be dependent on structure of input space X.
- Novikoff is only dependent on X through radius R.
- Expectation is that a typical ambient input space is only sparsely populated (cf manifold/cluster hypotheses).
- Pounce will depend on the cover of X.
- In particular the number of balls $\mathcal{N}(X, \rho)$ of diameter ρ.

Pounce Bound

Theorem

The mistakes M of POUNCE are bounded by

$$M \leq \mathcal{N}(X, \rho) + \|\boldsymbol{u}\|^2 \rho + 1 \quad ,$$

for all $0 < \rho$, and for all $\boldsymbol{u} \in \mathbb{R}^n$ such that

$$\boldsymbol{u}(i_t) y_t \geq 1$$

for all $t = 1, \ldots, \ell$.

- **Definition:** $\mathcal{N}(X, \rho)$ is the minimum number of balls of squared diameter ρ that cover X.