

MACHINE LEARNING

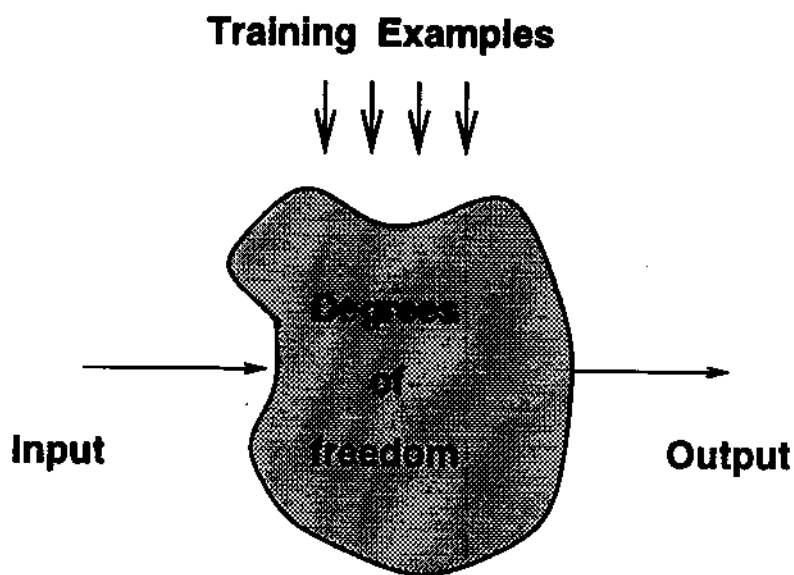
Yaser S. Abu-Mostafa

NSF Engineering Research Center

- **The learning problem**
- **Generalization**
- **Hints**
- **Distributed Learning**

Industry Day, May 17, 2001

LEARNING

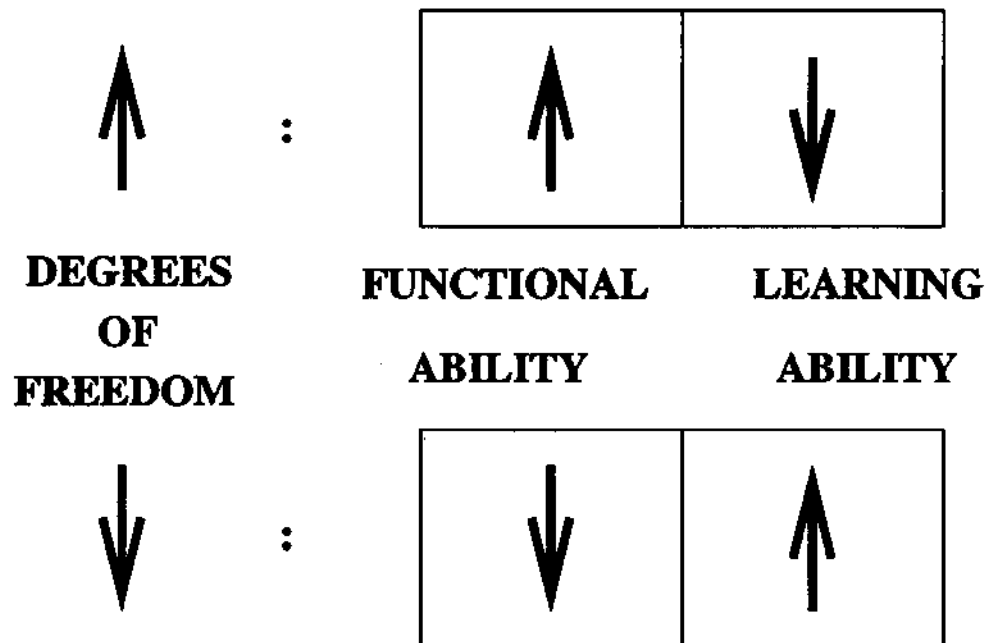


The Learning system can be:

GENERIC, such as a neural network

SPECIALIZED, such as a Vasicek model

DILEMMA



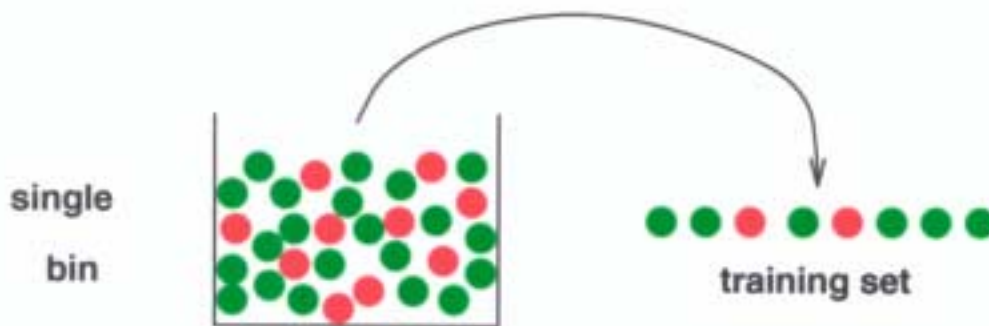
Flip 1 coin 10 times and obtain 10 heads \Rightarrow coin is probably not fair.

Flip 1000 coins 10 times each and find 1 coin which comes up heads all ten times. Are we confident that this coin is unfair?

If all 1000 coins are fair, the probability that at least 1 coin yields 10 heads is about 62%.

The Bin Model

Abstracts the relevant quantities from the target function and the learning model.



green marble : $g(x) = f(x)$
 red marble : $g(x) \neq f(x)$

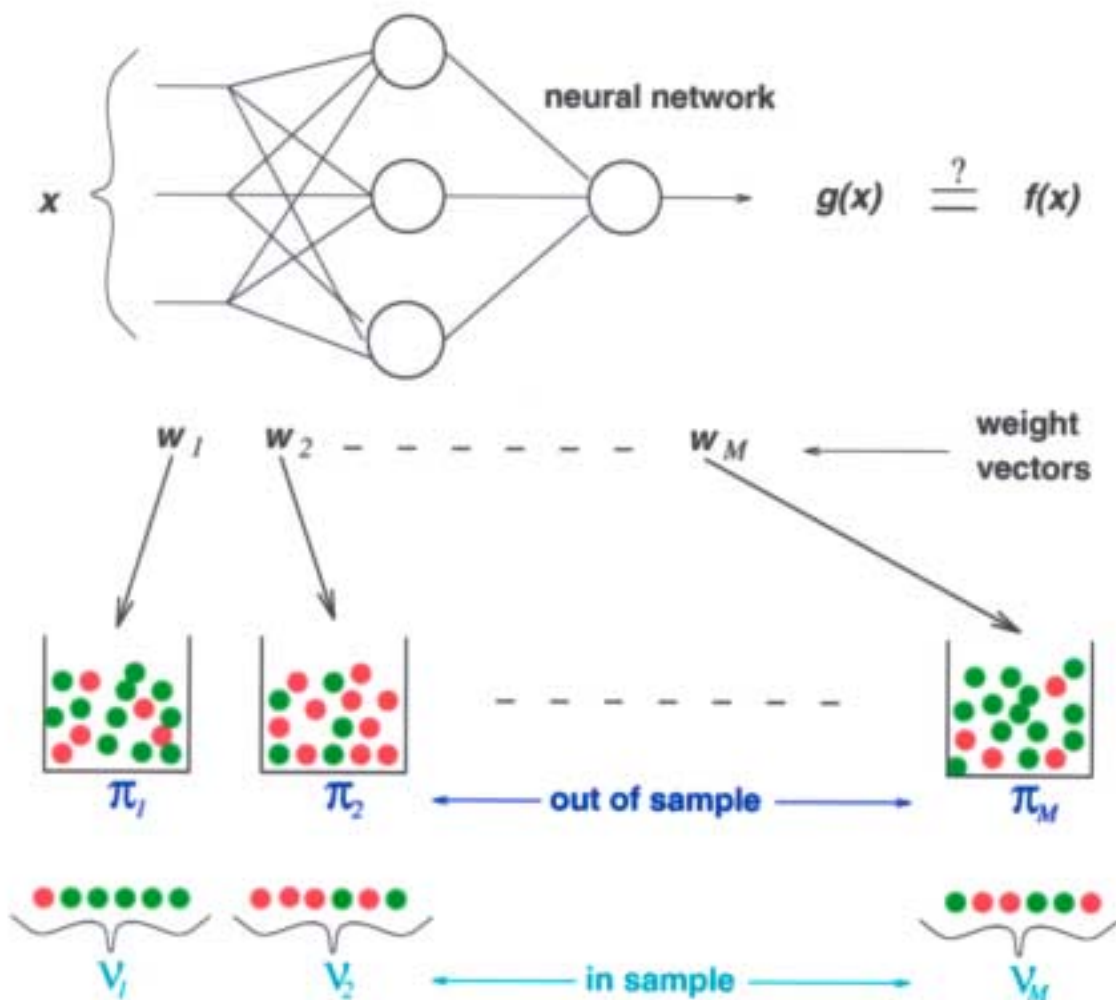
\swarrow hypothesis \searrow target

fraction of red marbles in the bin $\equiv \pi$ probability of error (out of sample)

fraction of red marbles in the training sample $\equiv v$ frequency of error (in sample)

Many Bins:

Any learning model (e.g., a neural network) can be represented by a set of bins.



HINTS

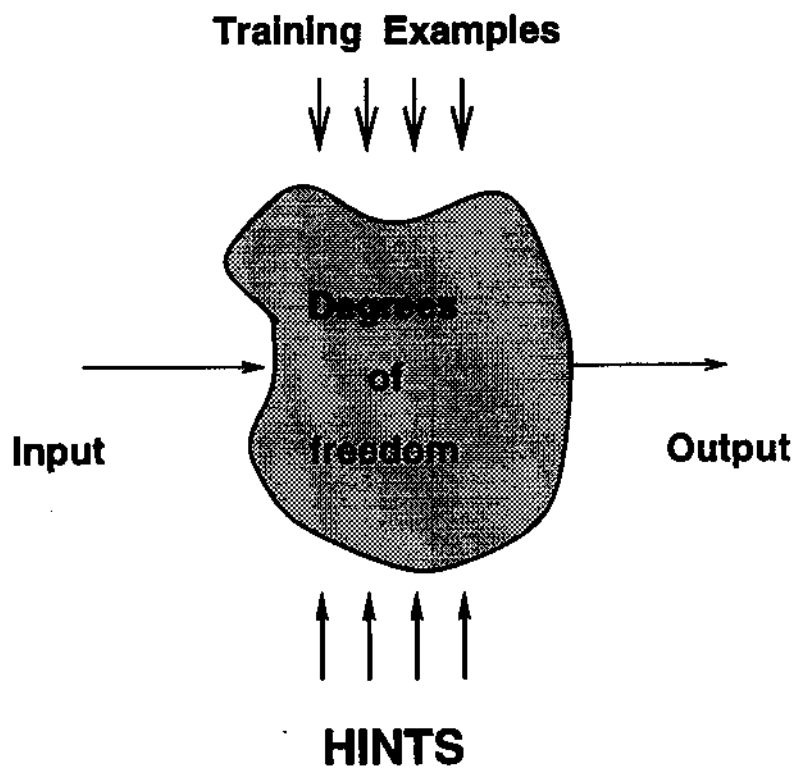
-- INFORMATION OTHER THAN THE DATA --

**COMMON
SENSE
RULES**



**SPECIAL
EXPERTISE**

LEARNING FROM HINTS



Hints are used as constraints on the learning process

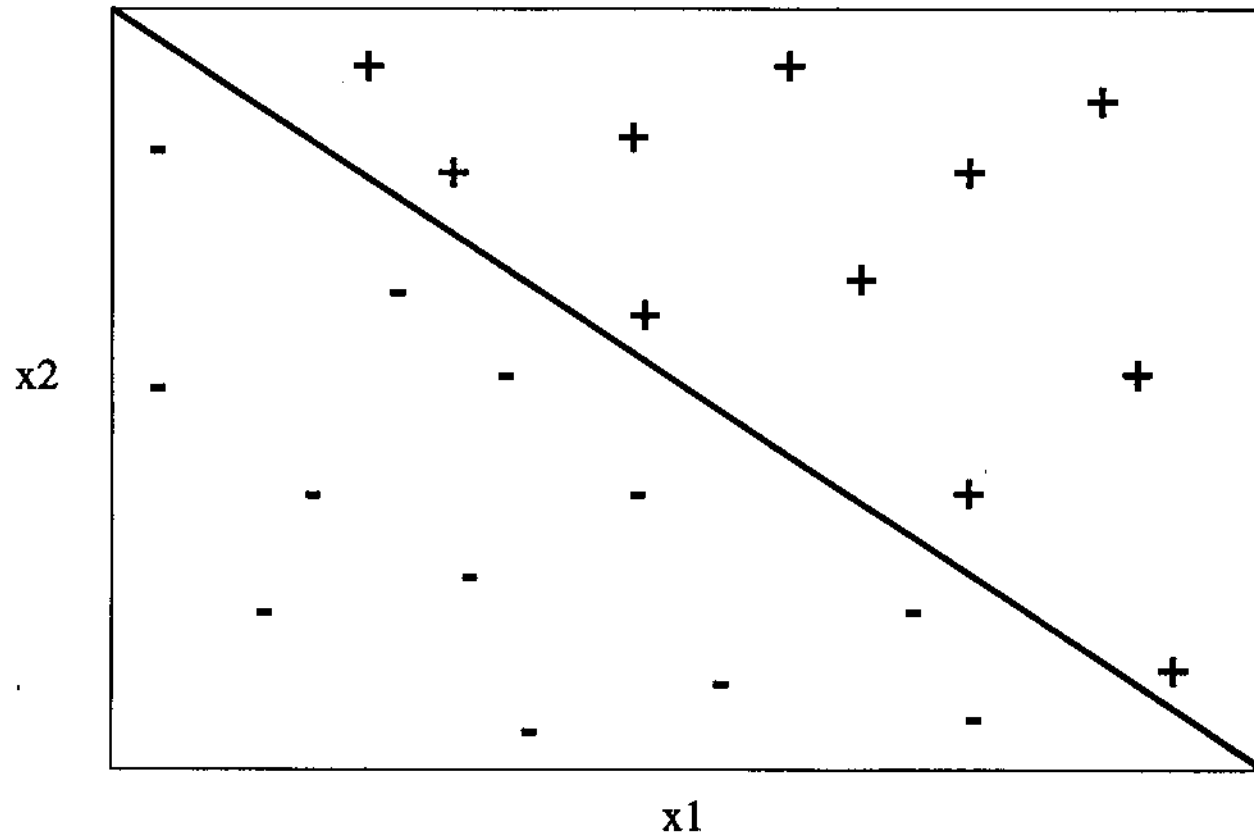
CREDIT HINT

Consider two applicants, applicant A and applicant B, who have identical profiles except that B has a higher salary

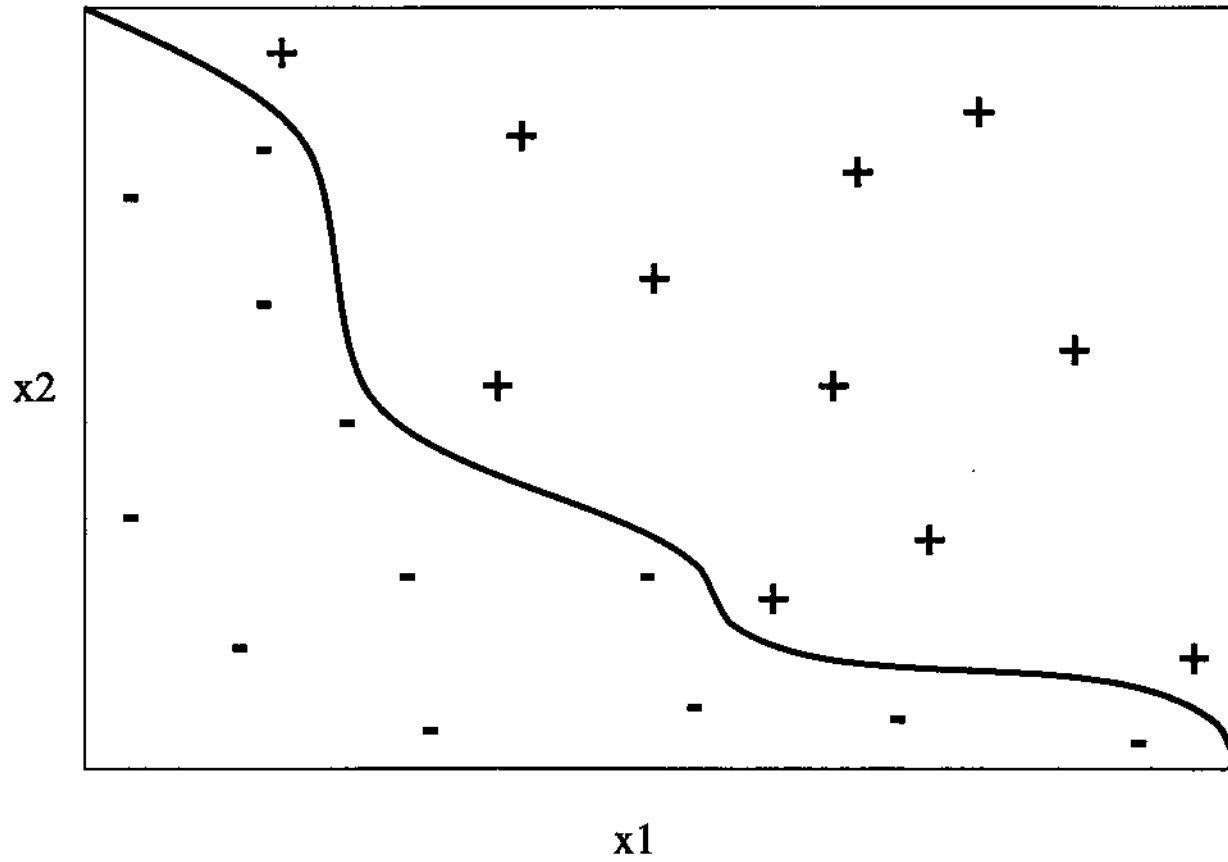
Feature	applicant A	applicant B
years at job	3	3
debt	\$5000	\$5000
salary	\$20,000	\$25,000
age	27	27

If A is approved, B should also be approved

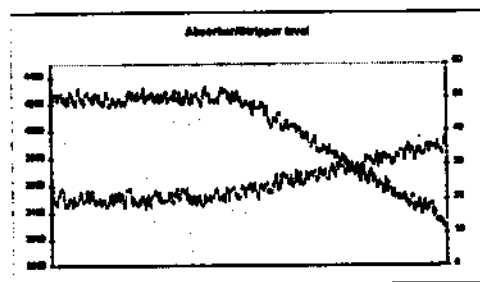
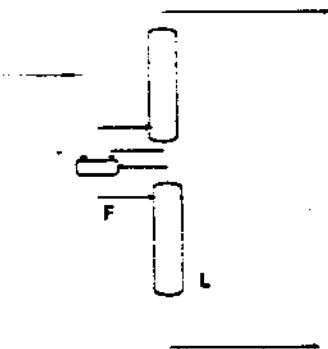
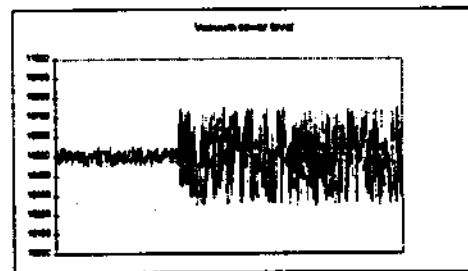
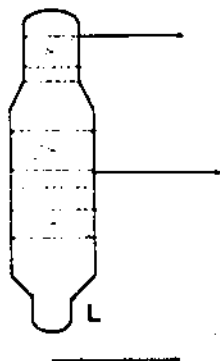
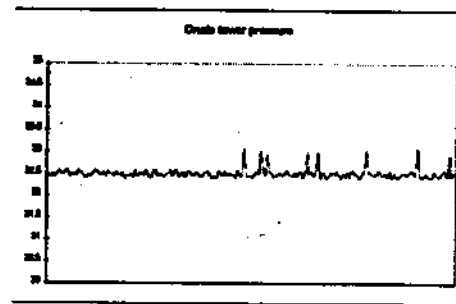
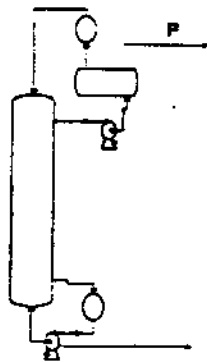
Linear Decision Boundary



Monotonic Decision Boundary



Signature recognition



FAULT MODES

- ANGULAR MISALIGNMENT
- PARALLEL MISALIGNMENT
- MOTOR BEARING MISALIGNMENT
- MOTOR BEARING WEAR
- MOTOR LAMINATION FAULT
- COUPLING WEAR

-
-
-
-
-



NORMAL



SLIGHT



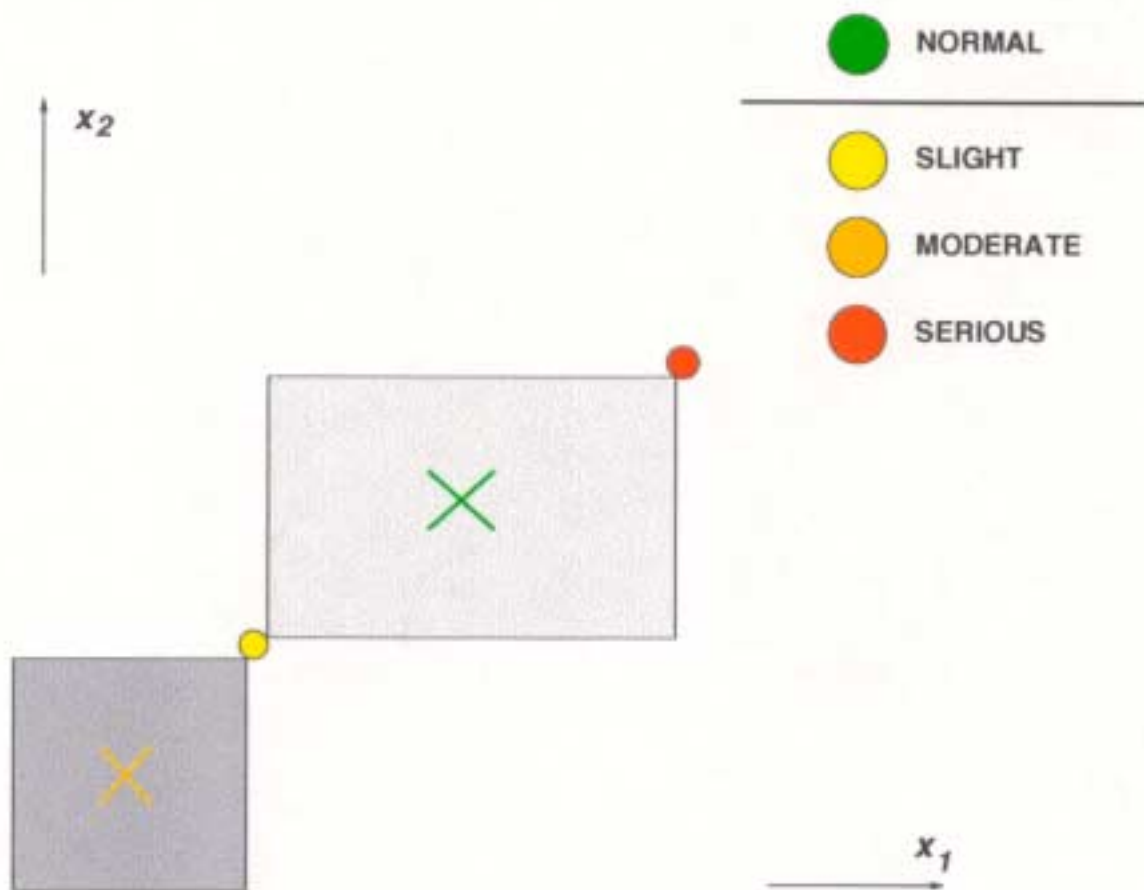
MODERATE



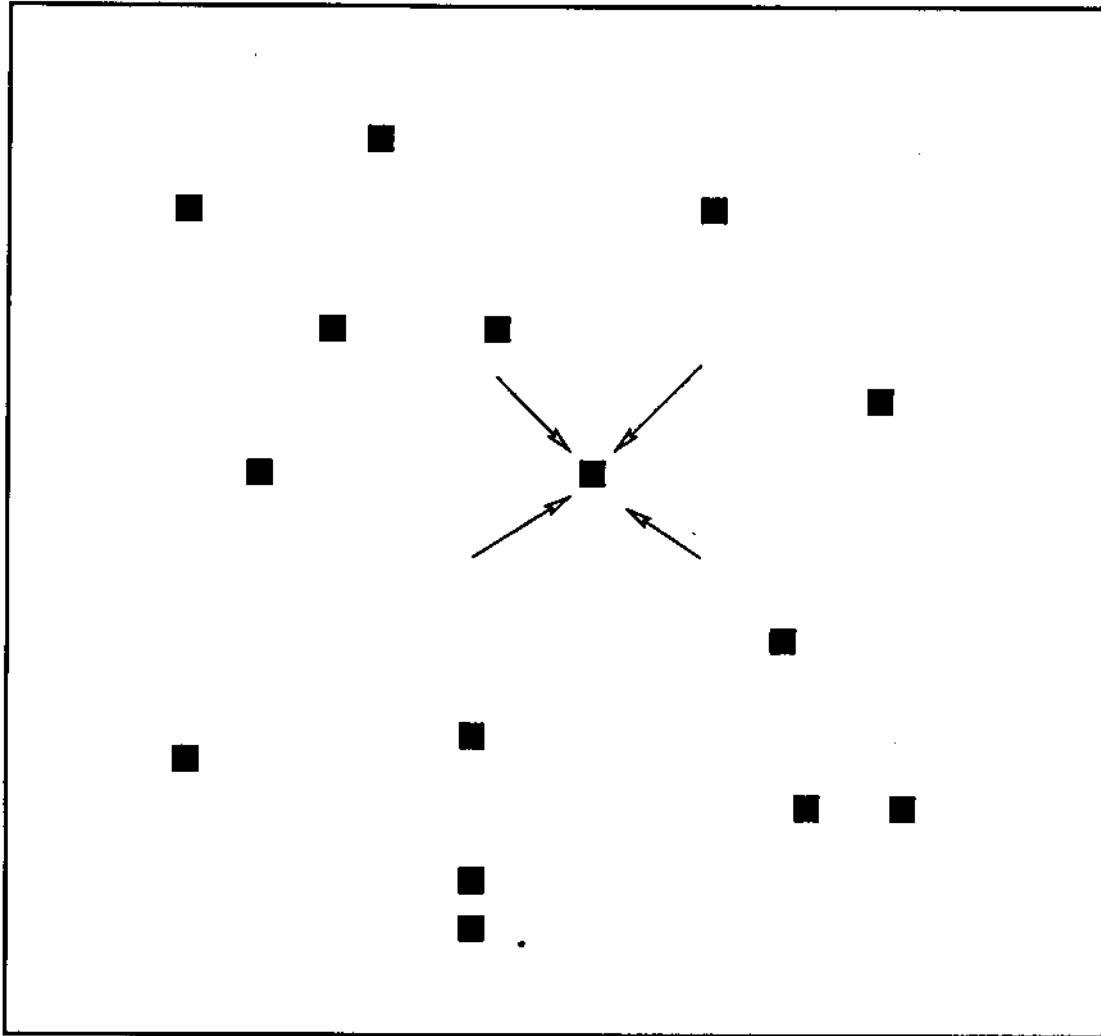
SERIOUS

MONOTONICITY HINT

The class centers are constrained by monotonicity in the input variables.

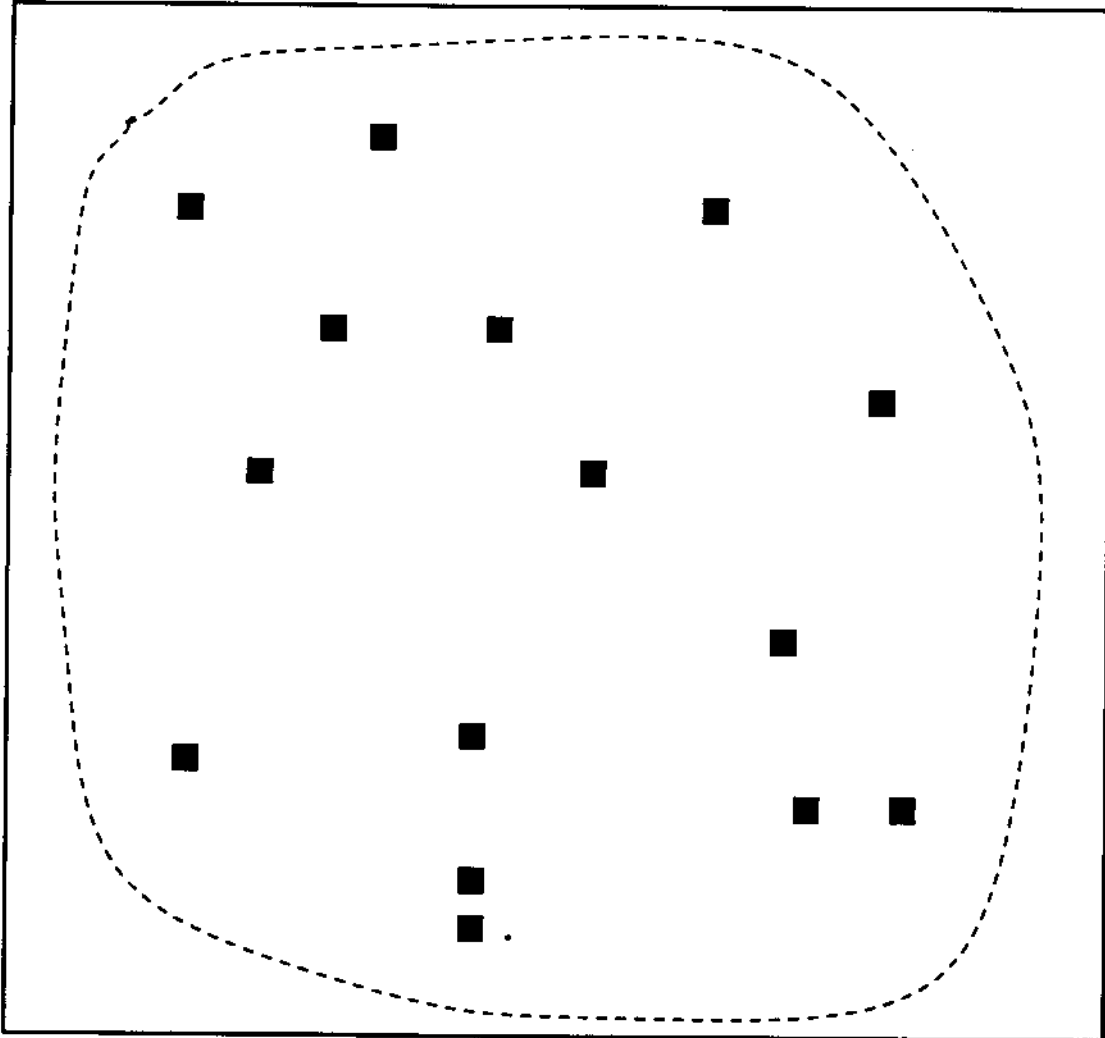


DISTRIBUTED LEARNING



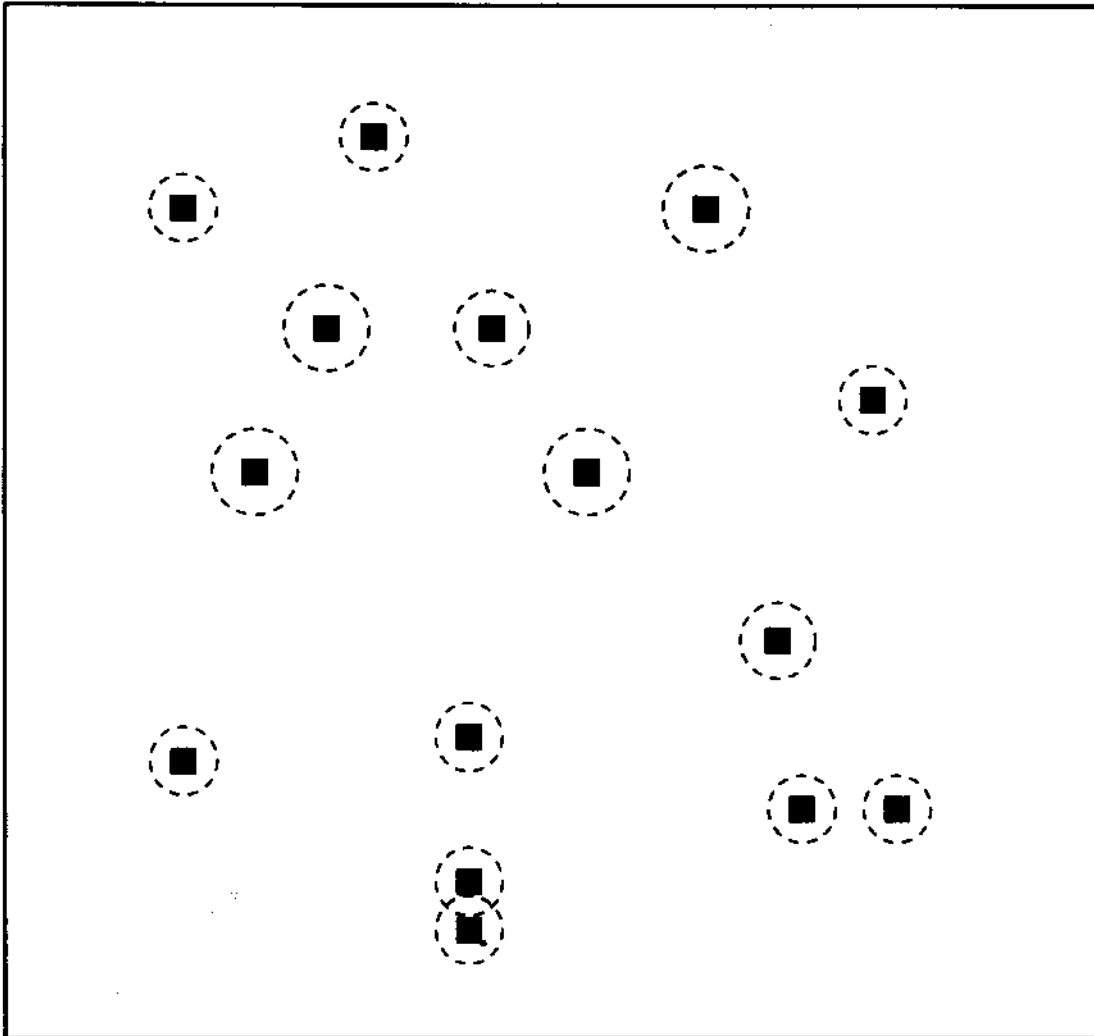
- LOCAL INPUTS
- LOCAL PARAMETERS

GLOBAL OBJECTIVE FUNCTION



- **Noise and slow convergence**
- **Not really 'distributed' learning**

LOCAL OBJECTIVE FUNCTION



- How to 'break up' the global objective?