Classification problems

- Exemplar characterised by a set of features; decide class to which exemplar belongs

Compare regression problems

- Exemplar characterised by a set of features;
- decide value of continuous output (*dependant) variable*

Classifying apples and pears

	Greeness	Height	Width	Taste	Weight	Height/Width	Class
No. 1	210	60	62	Sweet	186	0.97	Apple
No. 2	220	70	51	Sweet	180	1.37	Pear
No. 3	215	55	55	Tart	152	1.00	Apple
No. 4	180	76	40	Sweet	152	1.90	Pear
No. 5	220	68	45	Sweet	153	1.51	Pear
No. 6	160	65	68	Sour	221	0.96	Apple
No. 7	215	63	45	Sweet	140	1.40	Pear
No. 8	180	55	56	Sweet	154	0.98	Apple
No. 9	220	68	65	Tart	221	1.05	Apple
No. 10	190	60	58	Sour	174	1.03	Apple

To what class does this belong?

No. x	222	70	55	Sweet	185	1.27	?

k-Nearest Neighbour

D set of training samples

Find *k* nearest neighbours to **q** according to this similarity criterion

For each $\mathbf{x} \in \mathbf{X}$

$$Sim(\mathbf{q}, \mathbf{x}) = \sum_{f \in \mathbf{F}} w_f \cdot d(q_f, x_f)$$

where

$$d(q_f, x_f) = \begin{cases} 0 & f \text{ discrete and } q_f = x_f \\ 1 & f \text{ discrete and } q_f \neq x_f \\ \left| q_f - x_f \right| & f \text{ continuous} \end{cases}$$
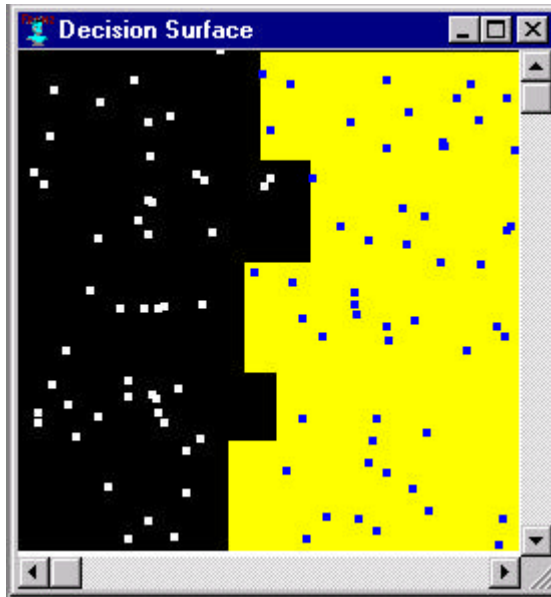
Condensed NN

D set of training samples

Find E where $E \subset D$; NN rule used with E
should be as good as D

choose $x \in D$ randomly,
$D \leftarrow D \setminus \{x\}$, $E \leftarrow \{x\}$,
DO
    learning? $\leftarrow$ FALSE,
   FOR EACH $x \in D$
       classify x by NN using E,
       if classification incorrect
           then $E \leftarrow E \cup \{x\}$,
                 $D \leftarrow D \setminus \{x\}$,
                 learning $\leftarrow$ TRUE,
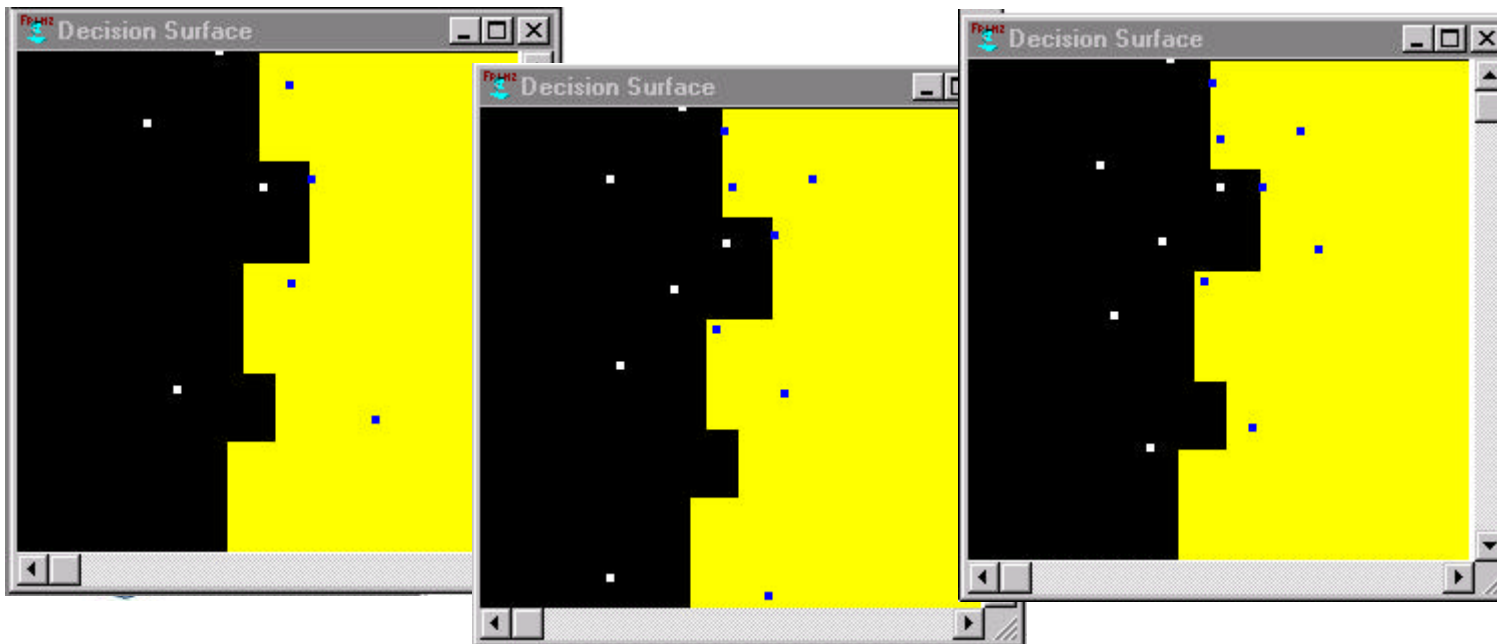WHILE (learning? $\neq$ FALSE)
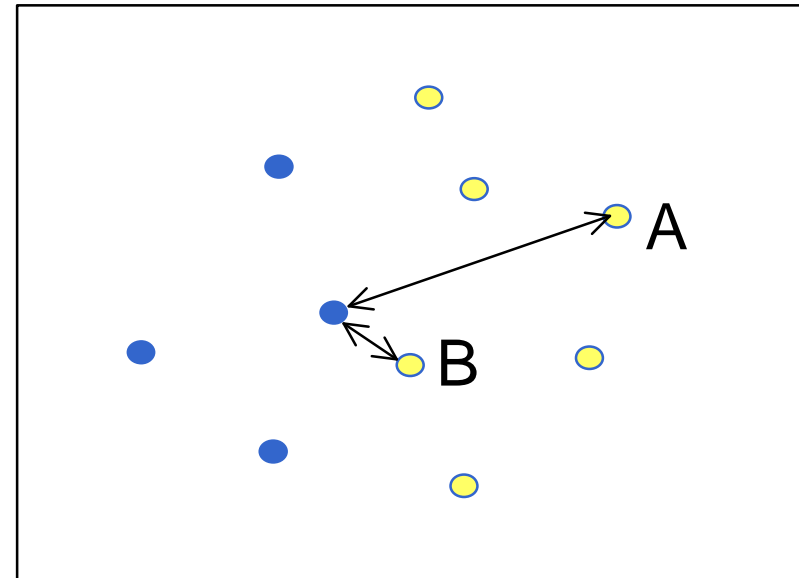
Condensed NN

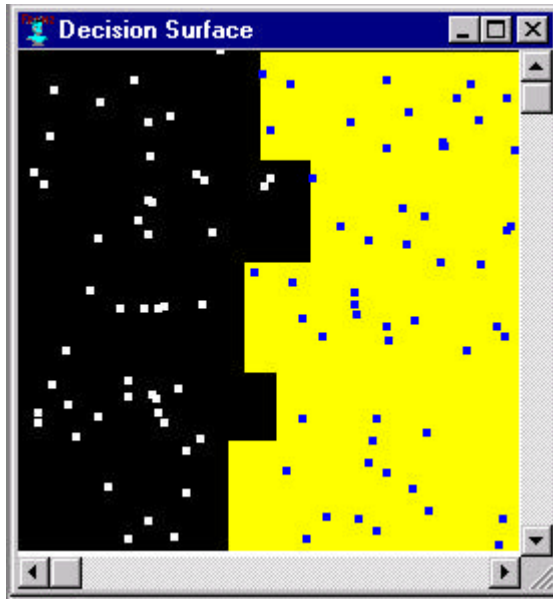100 examples
2 categories

Different CNN solutions

Improving Condensed NN

- Different outcomes depending on data order
  - that's a bad thing in an algorithm

- Sort data based on distance to nearest unlike neighbour
  - identify exemplars near decision surface
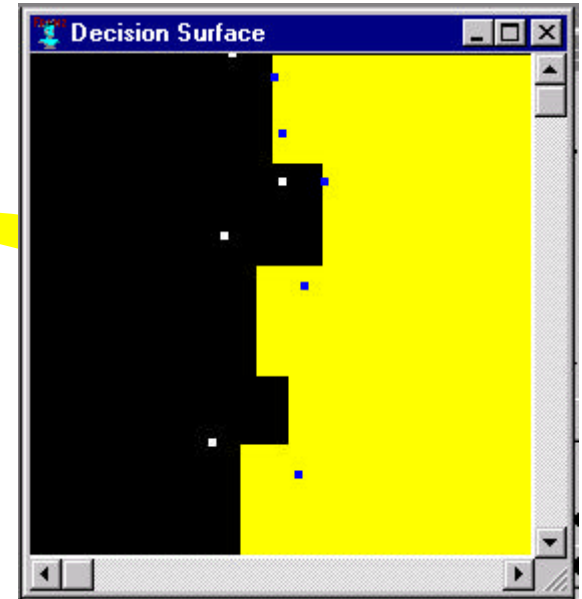 - in diagram

    B more useful than A

Condensed NN

100 examples
2 categories

CNN
using
NUN

Different CNN solutions