MVA based PID.

M. Babai

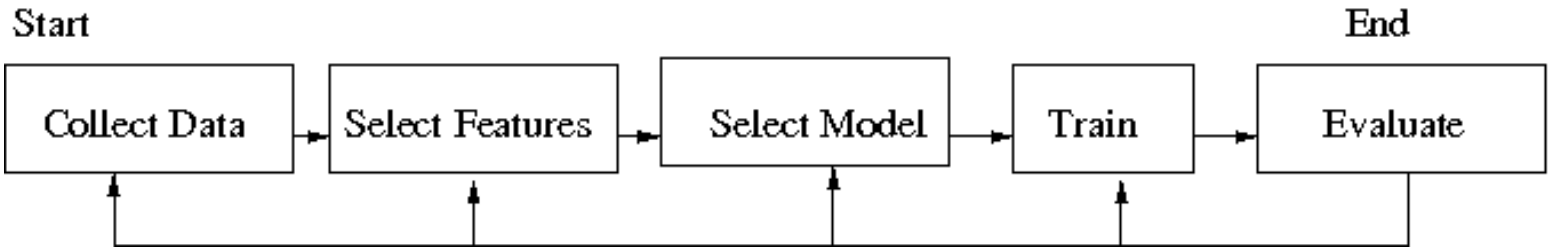
KVI / University of Groningen
The Netherlands.

July 25, 2012

M. Babai

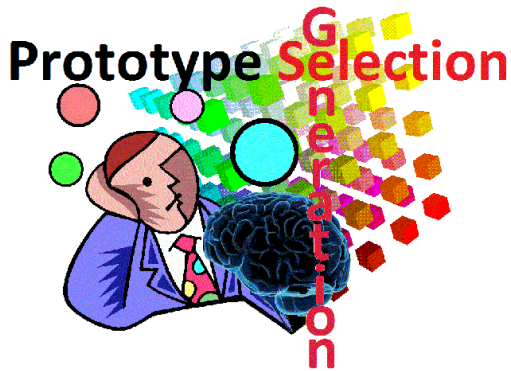
Outline

Design of a Pattern recognition system



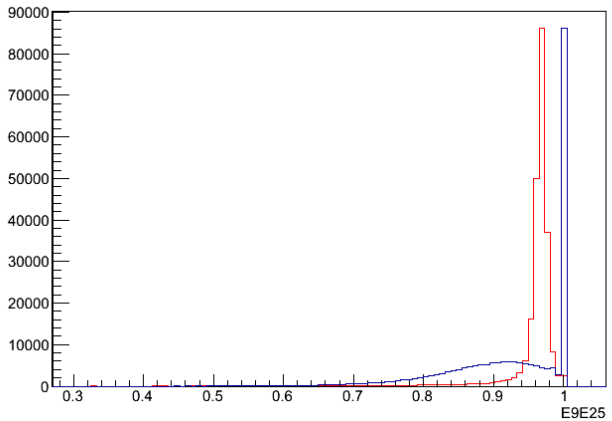
- Pre-processing: Select relevant information from data.
- Invariant: measurements do not have to change when the object appears in different context.
- Error-rate: Percentage of mis-assigned new patterns.
- Risk: Costs incorporation for each classification decision.
- (Cross-) Validation. Test set method, Leave one out, n-fold cross-validation, ...

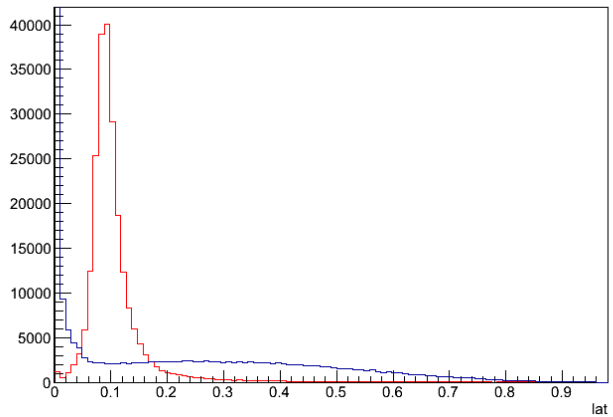
Classification challenges

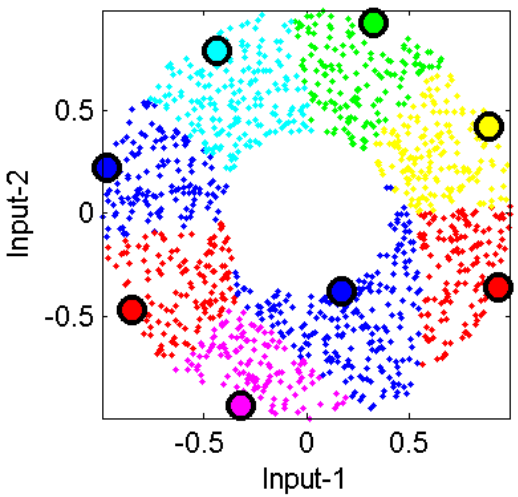


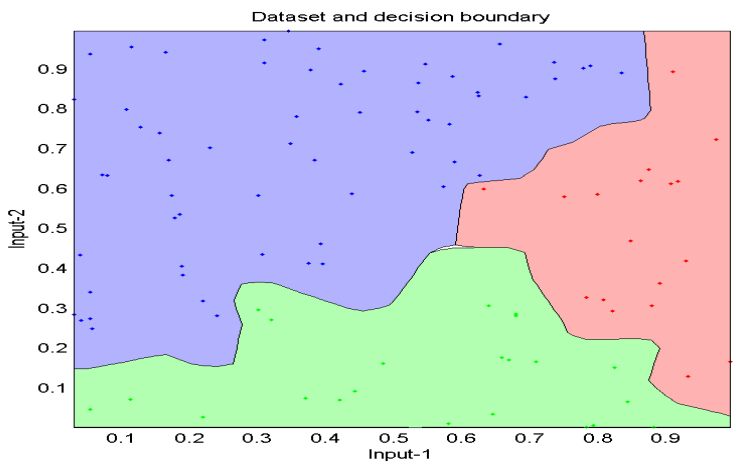
- The results depend on the variability of features.
- The variability can be affected by noise.
- How to cope with variability.

Motivation

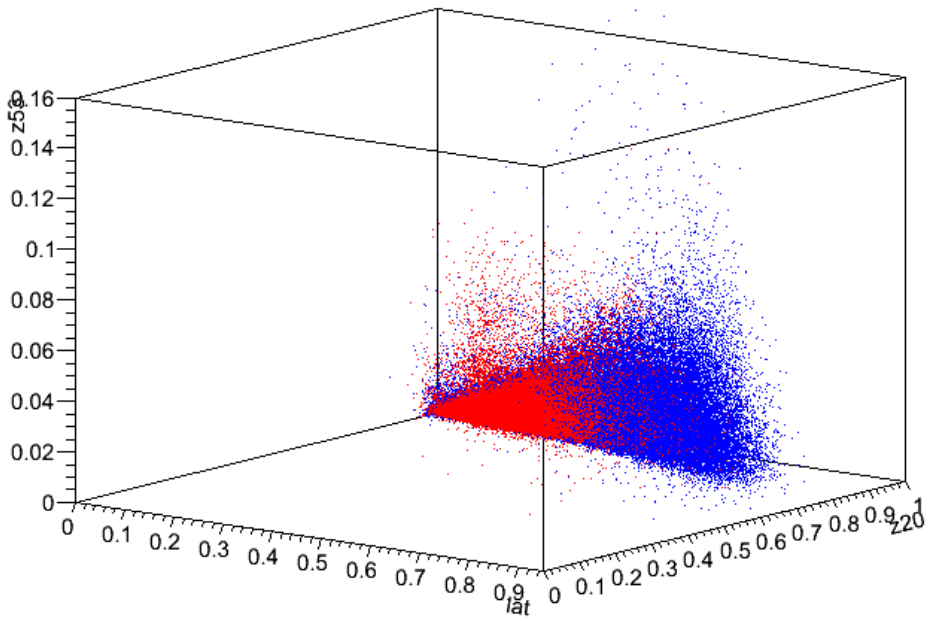

E9E25


lat


Input-1
Input-2


Dataset and decision boundary
Input-1
Input-2

Motivation


z5a
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
lat
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
z20

Available Algorithms.

- KNN (Density Estimator). Kd-tree based, standard (linear search).
  Pro: Easy to understand and use. Cons: Needs large data-set, relatively slow.
- Learning Vector Quantization (LVQ). LVQ1 and LVQ2.1 algorithms.
  Pros: Fast, small and easy to understand. Cons: Outputs are distances, difficult to find the optimal parameter set, time consuming training phase.

From TMVA (available in ROOT)

- Multilayer Perceptron (MLP).
- Boosted Decision Trees (BDT).

These two algorithms are just interfaced from TMVA.

Available Algorithms (*Pre-processing*).

- Principal Component Analysis (PCA) based parameter transformation.
- K-Means Clustering.
  Proto-type initialization. *"Un-supervised"* class mean based clustering. (Could also be used for winner takes all classification.)

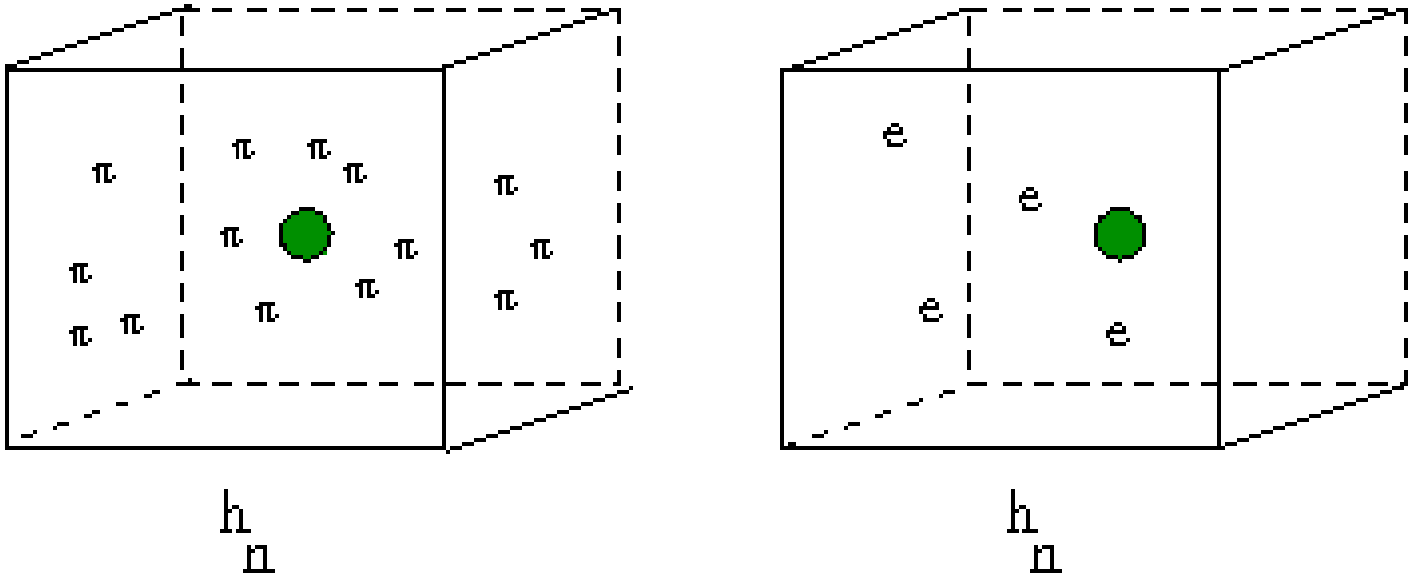
K- Nearest Neighbors

$$p_n(x) = \frac{K_n/n}{V_n}$$

The cell is expanded until it encompasses K_n samples.
a posteriori probability is merely the fraction of samples within a cell with the label, k_i/k.
The Bayes decision rule becomes:

$$P(\omega_j|x) = \underset{i}{max} P(\omega_i|x).$$

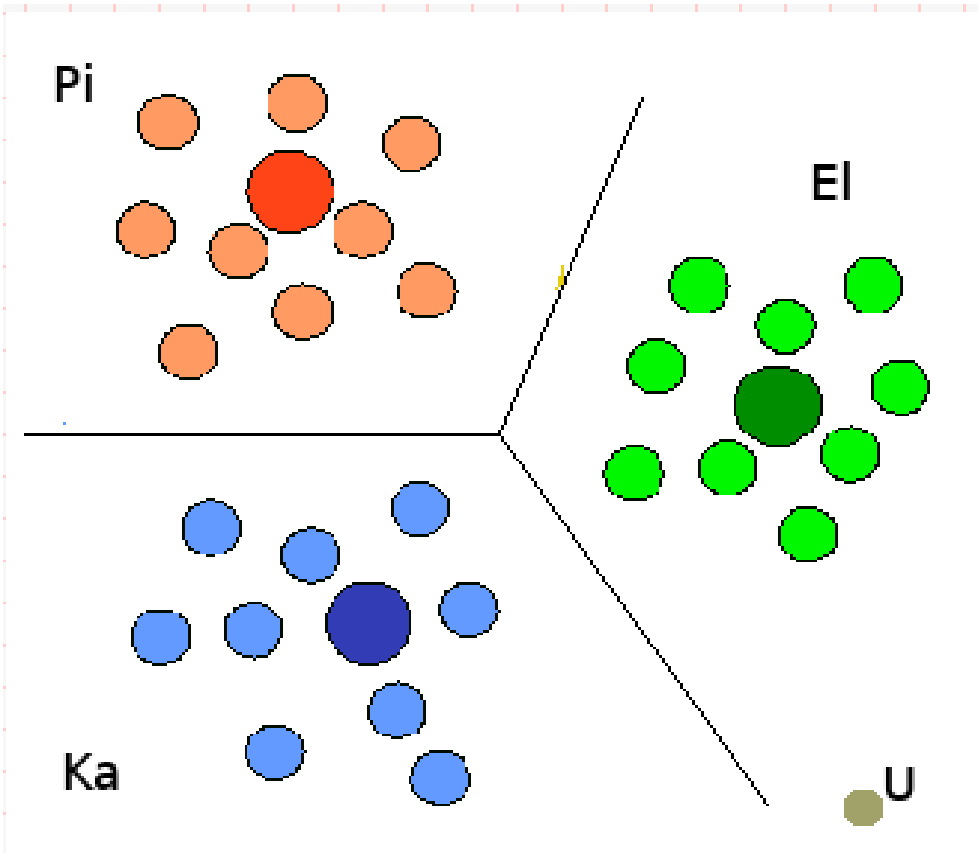
Probability densities (KNN)



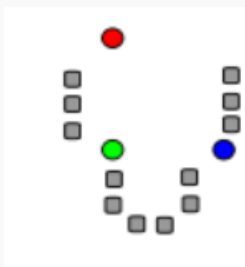
$$\pi \text{ and } e \in \{\text{Training data}\}$$

$$p_n(x) \propto \left(\frac{\#training\ elements}{h_n} \right)$$

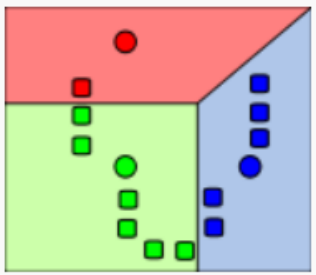
Learning Vector Quantization



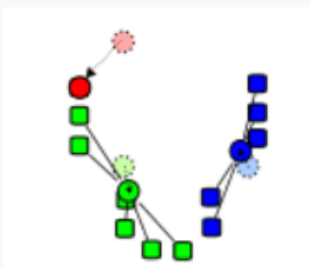
K-Means



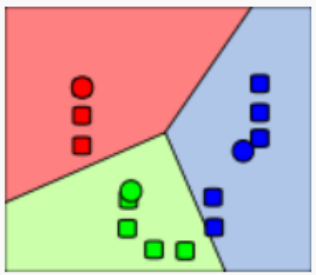
1) k initial "means" (in this case k=3) are randomly selected from the data set (shown in color).



2) k clusters are created by associating every observation with the nearest mean. The partitions here represent the Voronoi diagram generated by the means.

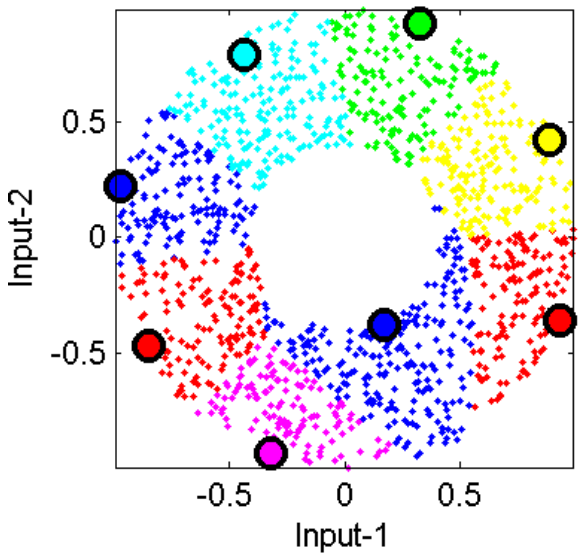


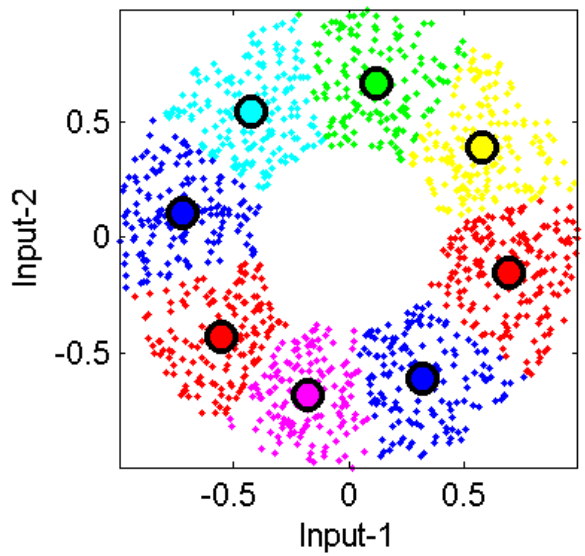
3) The centroid of each of the k clusters becomes the new mean.



4) Steps 2 and 3 are repeated until convergence has been reached.

k-Means example





Using Mva's in PANDA

- There are macro's and example programs available in "pandaroot/PndTools/MVA/".
- Documentation.
- Included parameters [E/$p_{(emc)}$, lat, z20, z53, E9E25].
- Labels $\{e^-, \pi^-\}$.

Questions?