

Cluster analysis, artificial neural networks and support vector machines in sensory analysis

E. Marchitto, R. De Leone, C. Elisei, M. Farina

University of Camerino, Italy

L. Bailetti

CIAS, Centro Italiano di Analisi Sensoriale, Matelica (MC), Italy

Outline of this talk

The purpose of this talk is to discuss the issues associated with selection of assessors for sensory panels. Usually, the specialist (Panel Leader) based on his/her subjective experience selects assessors able to discern the characteristics of the organoleptic properties of food.

- ☞ To evaluate and select those assessors we used 3 different mathematical approaches:
 - ⇒ Cluster Analysis;
 - ⇒ Artificial Neural Networks;
 - ⇒ Support Vector Machines.
- ☞ Computational results;
- ☞ Conclusions and future directions.

Outline of this talk

Problem Description

Cluster Analysis

Artificial Neural Networks

Support Vector Machines

Computational results

Conclusions

Problem Description

Outline of this talk

Problem Description

Sensory Analysis, 1

Sensory Analysis, 2

Sensory Analysis, 3

Sensory Analysis, 4

Cluster Analysis

Artificial Neural
Networks

Support Vector
Machines

Computational
results

Conclusions

Sensory analysis is a well-known method in food-industry to characterize organoleptic properties of different kinds of food.

To carry out the measurement of organoleptic sensations the method of sensory analysis exploits the psychology of human beings.

Sensory analysis provides information on specific products. Samples consist in food and the instruments are the tasters.

Outline of this talk

Problem Description

Sensory Analysis, 1

Sensory Analysis, 2

Sensory Analysis, 3

Sensory Analysis, 4

Cluster Analysis

Artificial Neural Networks

Support Vector Machines

Computational results

Conclusions

The gustatory and olfactory properties in the investigated food arise as perception through the tasters consciousness and he/she has to express them. Therefore, in the sensory profile obtained by sensory analysis, the sensitivity of the taster is strictly linked to the organoleptic properties of food.

Still, we can analyze the subjective results and obtain valid conclusions applying statistical instruments.

Descriptive analysis involves trained tasters who evaluate products by rating the intensity of various characteristics on a scale. Different statistical analysis are applied to some characteristics of interest to find out differences among various products.

Outline of this talk

Problem Description

Sensory Analysis, 1

Sensory Analysis, 2

Sensory Analysis, 3

Sensory Analysis, 4

Cluster Analysis

Artificial Neural
Networks

Support Vector
Machines

Computational
results

Conclusions

The **Panel Leader** is the coordinator of the tasters' team and guides and controls all the operations of the sensory analysis.

He/she has a specific scientific training deriving also from different technical and scientific disciplines (such as chemistry, mathematics, statistics and physics) that can be used as support to the sensory analysis, as well as good organization capacities.

Outline of this talk

Problem Description

Sensory Analysis, 1

Sensory Analysis, 2

Sensory Analysis, 3

Sensory Analysis, 4

Cluster Analysis

Artificial Neural
Networks

Support Vector
Machines

Computational
results

Conclusions

The **Panel** is a group of people (or assessors) trained for a number of evaluations, and is the main instrument of the sensory analysis.

These people, guided by the Panel Leader, can be inexperienced assessors (consumers) or trained assessors. This Panel is created after a selection and training phase that meets **UNI-ISO 8586-1** and **UNI-ISO 8586-2**.

[ISO 8586-1. Sensory Analysis - General guidance for the selection, training and monitoring of assessors-Part 1: Selected Assessors, 1993].

[ISO 8586-2. Sensory Analysis - General guidance for the selection, training and monitoring of assessors-Part 2: Expert, 1994].

Outline of this talk

Problem Description

Sensory Analysis, 1

Sensory Analysis, 2

Sensory Analysis, 3

Sensory Analysis, 4

Cluster Analysis

Artificial Neural Networks

Support Vector Machines

Computational results

Conclusions

Cluster Analysis

Outline of this talk

Problem Description

Cluster Analysis

Cluster Analysis, 1

Cluster Analysis, 2

Artificial Neural
Networks

Support Vector
Machines

Computational
results

Conclusions

Cluster Analysis allows to partition elements into groups such that:

- ★ elements in the same cluster have a high degree of similarity (share some common traits);
- ★ there is a weak degree of association among elements in different clusters.

[Outline of this talk](#)

[Problem Description](#)

[Cluster Analysis](#)

[Cluster Analysis, 1](#)

[Cluster Analysis, 2](#)

[Artificial Neural Networks](#)

[Support Vector Machines](#)

[Computational results](#)

[Conclusions](#)

The main ingredients in Cluster Analysis are the choice of the algorithm, the distance function between objects when forming the clusters and amalgamation/linkage rules.

- **Algorithms:** there are different categories of algorithms for clustering; in our experiments we used both Hierarchical Clustering Methods that create a hierarchical decomposition of the dataset with successive fusions or finer separations and Partition Algorithms that decompose the dataset into distinct nonoverlapping groups, evaluate the quality of the partition and redefine the decomposition;

Outline of this talk

Problem Description

Cluster Analysis

Cluster Analysis, 1

Cluster Analysis, 2

Artificial Neural Networks

Support Vector Machines

Computational results

Conclusions

The main ingredients in Cluster Analysis are the choice of the algorithm, the distance function between objects when forming the clusters and amalgamation/linkage rules.

- **Similarity/Dissimilarity Measure:** several distinct distances can be defined for each pair of objects; in our experiments we obtained the best rules when using the Manhattan's correlation measure;
- **Amalgamation/Linkage Rules:** different methods can be also used to evaluate the distance between clusters; in our experiments we obtained the best results when using the Complete's linkage.

For this experiment we used the statistical software Minitab.

Outline of this talk

Problem Description

Cluster Analysis

Cluster Analysis, 1

Cluster Analysis, 2

Artificial Neural Networks

Support Vector Machines

Computational results

Conclusions

Artificial Neural Networks

Outline of this talk

Problem Description

Cluster Analysis

**Artificial Neural
Networks**

Artificial Neural
Network Model:
The Multilayer
Perceptron
Structure of a MLP
Network
Characteristics of
MLP

Support Vector
Machines

Computational
results

Conclusions

Artificial Neural Network Model: The Multilayer Perceptron

Important aspects in the Artificial Neural Networks are the definition of the number of nodes and links in the network (network topology), the definition of training and testing sets and the training algorithm.

In our experiments we used a Multilayer Perceptron.

- ✧ Units are organized in levels/layers (input level, output level, one or more hidden levels);
- ✧ Units in each level are fully connected only to the units of the next level;
- ✧ Each unit has a single output;
- ✧ The training procedure computes the weights on the arcs.

Outline of this talk

Problem Description

Cluster Analysis

Artificial Neural
Networks
Artificial Neural
Network Model:
The Multilayer
Perceptron

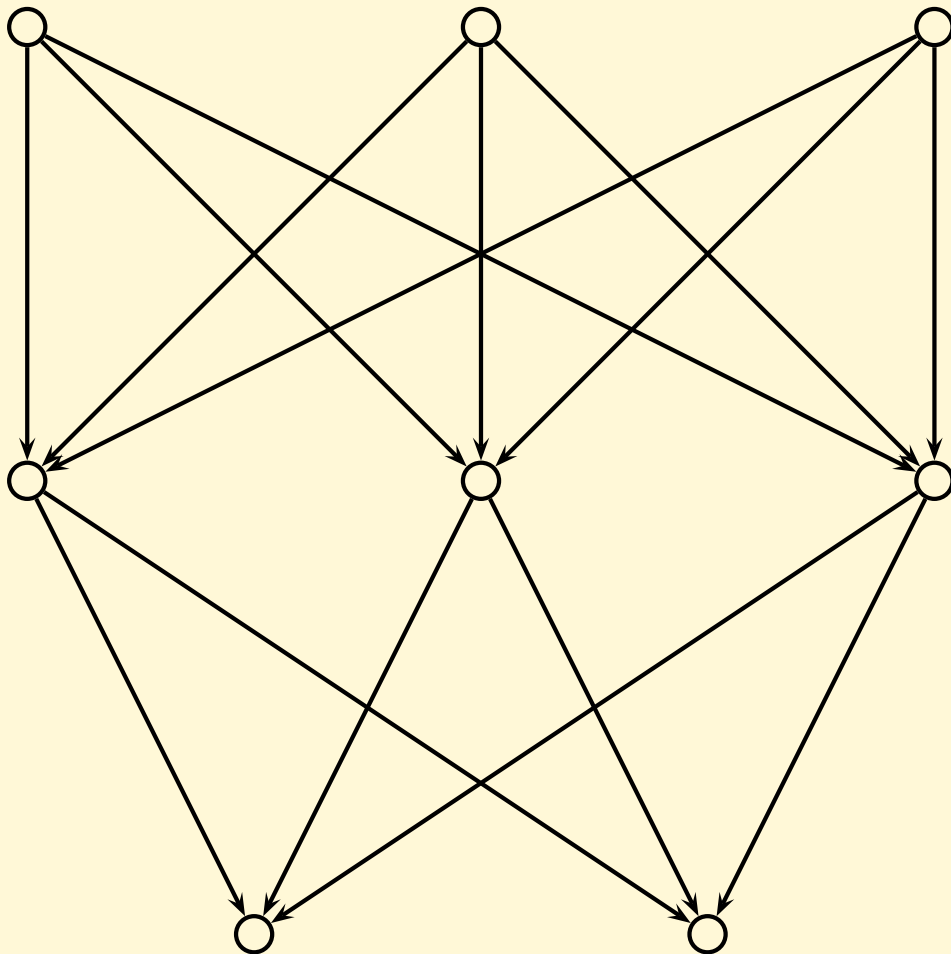
Structure of a MLP
Network
Characteristics of
MLP

Support Vector
Machines

Computational
results

Conclusions

Structure of a MLP Network



input layer

hidden layer

output layer

[Outline of this talk](#)

[Problem Description](#)

[Cluster Analysis](#)

Artificial Neural Networks

Artificial Neural Network Model: The Multilayer Perceptron

Structure of a MLP Network

Characteristics of MLP

[Support Vector Machines](#)

[Computational results](#)

[Conclusions](#)

Characteristics of MLP

- ✦ Composed by several perceptrons connected in cascade;
- ✦ The main problem is the training (modification of weights in the hidden layers);
- ✦ Number of hidden layers;
- ✦ The Back-propagation procedure (and its variants) is one of the most used methods for the training procedure;
- ✦ Structure.

Outline of this talk

Problem Description

Cluster Analysis

Artificial Neural Networks

Artificial Neural Network Model:
The Multilayer Perceptron

Structure of a MLP Network

Characteristics of MLP

Support Vector Machines

Computational results

Conclusions

Support Vector Machines

Outline of this talk

Problem Description

Cluster Analysis

Artificial Neural
Networks

**Support Vector
Machines**

The Classification
Problem, 1

The Classification
Problem, 2

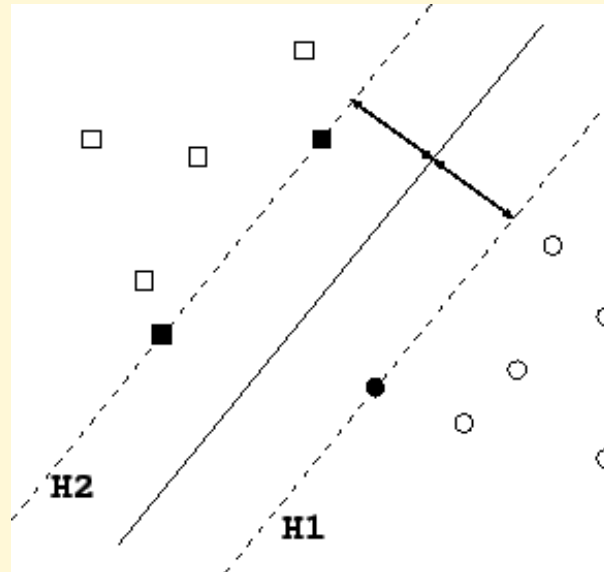
Dual Problem

Computational
results

Conclusions

The Classification Problem, 1

Support Vector Machines are learning machines that can perform binary classification and real valued function approximation. For the linearly separable case, SVM calculates the maximum margin hyperplane separating the two sets of data.



Outline of this talk

Problem Description

Cluster Analysis

Artificial Neural Networks

Support Vector Machines
The Classification Problem, 1

The Classification Problem, 2

Dual Problem

Computational results

Conclusions

The Classification Problem, 2

A classification task usually involves a training set containing “target values” and several “attributes”. The goal of SVM is to produce a model that predict target values for new data instances.

$$\text{Training set } \mathcal{T} = \{(x^1, y_1), (x^2, y_2), \dots, (x^N, y_N)\}$$

$$\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

$$\begin{aligned} \min_{w, b, \xi} \quad & \frac{1}{2} w^T w + C \sum_{i=1}^N \xi_i \\ \text{subject to} \quad & y_i (w^T \phi(x^i) + b) \geq 1 - \xi_i, \quad i = 1, \dots, N \\ & \xi_i \geq 0, \quad i = 1, \dots, N \end{aligned}$$

Outline of this talk

Problem Description

Cluster Analysis

Artificial Neural Networks

Support Vector Machines

The Classification Problem, 1

The Classification Problem, 2

Dual Problem

Computational results

Conclusions

$$\begin{aligned} \max_{w,b} \quad & \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \alpha_i \alpha_j Q_{ij} \\ \text{subject to} \quad & \sum_{i=1}^N \alpha_i y_i = 0 \\ & 0 \leq \alpha_i \leq C, \quad i = 1, \dots, N \end{aligned}$$

Outline of this talk

Problem Description

Cluster Analysis

Artificial Neural Networks

Support Vector Machines

The Classification Problem, 1

The Classification Problem, 2

Dual Problem

Computational results

Conclusions

$$\begin{aligned} \max_{w,b} \quad & \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \alpha_i \alpha_j Q_{ij} \\ \text{subject to} \quad & \sum_{i=1}^N \alpha_i y_i = 0 \\ & 0 \leq \alpha_i \leq C, \quad i = 1, \dots, N \end{aligned}$$

where

$$Q_{ij} = y_i y_j K_{ij}$$

$$K_{ij} = K(x^i, x^j) = \phi(x^i)^T \phi(x^j)$$

$$K : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \quad \text{KERNEL}$$

Outline of this talk

Problem Description

Cluster Analysis

Artificial Neural Networks

Support Vector Machines

The Classification Problem, 1

The Classification Problem, 2

Dual Problem

Computational results

Conclusions

$$\begin{aligned} \max_{w,b} \quad & \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \alpha_i \alpha_j Q_{ij} \\ \text{subject to} \quad & \sum_{i=1}^N \alpha_i y_i = 0 \\ & 0 \leq \alpha_i \leq C, \quad i = 1, \dots, N \end{aligned}$$

The classification of new data x is based on the function

$$f(x) = \text{sign} \left(\sum_{i=1}^N y_i \alpha_i^* K(x^i, x) + b^* \right)$$

Outline of this talk

Problem Description

Cluster Analysis

Artificial Neural Networks

Support Vector Machines

The Classification Problem, 1

The Classification Problem, 2

Dual Problem

Computational results

Conclusions

Computational results

Outline of this talk

Problem Description

Cluster Analysis

Artificial Neural
Networks

Support Vector
Machines

**Computational
results**

Computational
results, 1

Computational
results, 2

Computational
results, 3

Computational
results, 4

Computational
results, 5

Conclusions

Computational results, 1

Different clustering methods have been tested to classify a dataset of 100 different assessors into 2 or 3 distinct groups:

- ➡ APPROVED/REJECTED;
- ➡ APPROVED/NON-DECISION/REJECTED.

In the first case, we added to a dataset a control set made of 6 additional data, 3 for each class.

In the second case, we still added 6 additional data to the dataset, but in this case only 2 for each of the 3 classes.

Outline of this talk

Problem Description

Cluster Analysis

Artificial Neural Networks

Support Vector Machines

Computational results
Computational results, 1

Computational results, 2

Computational results, 3

Computational results, 4

Computational results, 5

Conclusions

Computational results, 2

Among all the different possible choices of algorithms, distances and linkage rules, we concentrated our attention on the cases where the control sets have been correctly clustered.

The best results have been obtained when using a hierarchical method with Complete's Linkage method and a distance given by the Manhattan's correlation.

classes	APPROVED/REJECTED respect to Panel Leader
2	78.30%
3	88.31%

Outline of this talk

Problem Description

Cluster Analysis

Artificial Neural
Networks

Support Vector
Machines

**Computational
results**

Computational
results, 1

**Computational
results, 2**

Computational
results, 3

Computational
results, 4

Computational
results, 5

Conclusions

The same dataset has been analyzed using MLP:

- ✓ 106 examples (assessors);
- ✓ 9 inputs (features);
- ✓ 2 outputs (APPROVED/REJECTED);
- ✓ 1 hidden layer with 5 units;
- ✓ 90 connections (full connection).

The entire dataset has been utilized as learning set using the Panel Leader's proposal as output, where (1,0) corresponds to APPROVED, and (0,1) to REJECTED. The assessors are correctly classified when we use the standard Backpropagation algorithm in the learning phase.

Outline of this talk

Problem Description

Cluster Analysis

Artificial Neural Networks

Support Vector Machines

Computational results

Computational results, 1

Computational results, 2

Computational results, 3

Computational results, 4

Computational results, 5

Conclusions

Computational results, 4

	Panel Leader	Complete Manhattan (3 cl)	K-Means (2 cl)	Neural Networks (2 cl)
Assessors 1	REJECTED	NO DECISION	APPROVED	REJECTED
Assessors 2	REJECTED	REJECTED	APPROVED	REJECTED
Assessors 3	APPROVED	K-Means (2 cl)	APPROVED	APPROVED
Assessors 4	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 5	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 6	APPROVED	NO DECISION	APPROVED	APPROVED
Assessors 7	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 8	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 9	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 10	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 11	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 12	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 13	REJECTED	APPROVED	APPROVED	REJECTED
Assessors 14	REJECTED	REJECTED	REJECTED	REJECTED
Assessors 15	REJECTED	APPROVED	APPROVED	REJECTED
Assessors 16	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 17	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 18	REJECTED	NO DECISION	APPROVED	REJECTED
Assessors 19	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 20	REJECTED	REJECTED	REJECTED	REJECTED
Assessors 21	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 22	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 23	REJECTED	APPROVED	APPROVED	REJECTED
Assessors 24	REJECTED	APPROVED	APPROVED	REJECTED
Assessors 25	REJECTED	NO DECISION	APPROVED	REJECTED
Assessors 26	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 27	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 28	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 29	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 30	APPROVED	NO DECISION	APPROVED	APPROVED
Assessors 31	APPROVED	NO DECISION	APPROVED	APPROVED
Assessors 32	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 33	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 34	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 35	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 36	APPROVED	NO DECISION	APPROVED	APPROVED
Assessors 37	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 38	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 39	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 40	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 41	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 42	REJECTED	APPROVED	APPROVED	REJECTED
Assessors 43	REJECTED	APPROVED	APPROVED	REJECTED
Assessors 44	APPROVED	REJECTED	REJECTED	APPROVED
Assessors 45	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 46	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 47	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 48	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 49	REJECTED	APPROVED	APPROVED	REJECTED
Assessors 50	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 51	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 52	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 53	APPROVED	NO DECISION	APPROVED	APPROVED

	Panel Leader	Complete Manhattan (3 cl)	K-Means (2 cl)	Neural Networks (2 cl)
Assessors 54	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 55	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 56	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 57	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 58	APPROVED	NO DECISION	APPROVED	APPROVED
Assessors 59	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 60	APPROVED	NO DECISION	APPROVED	APPROVED
Assessors 61	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 62	APPROVED	NO DECISION	APPROVED	APPROVED
Assessors 63	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 64	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 65	REJECTED	NO DECISION	APPROVED	REJECTED
Assessors 66	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 67	REJECTED	REJECTED	APPROVED	REJECTED
Assessors 68	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 69	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 70	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 71	REJECTED	APPROVED	APPROVED	REJECTED
Assessors 72	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 73	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 74	APPROVED	NO DECISION	APPROVED	APPROVED
Assessors 75	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 76	REJECTED	APPROVED	APPROVED	REJECTED
Assessors 77	REJECTED	APPROVED	APPROVED	REJECTED
Assessors 78	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 79	REJECTED	APPROVED	APPROVED	REJECTED
Assessors 80	REJECTED	REJECTED	APPROVED	REJECTED
Assessors 81	REJECTED	APPROVED	APPROVED	REJECTED
Assessors 82	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 83	REJECTED	NO DECISION	APPROVED	REJECTED
Assessors 84	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 85	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 86	APPROVED	REJECTED	APPROVED	APPROVED
Assessors 87	APPROVED	NO DECISION	APPROVED	APPROVED
Assessors 88	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 89	REJECTED	NO DECISION	APPROVED	APPROVED
Assessors 90	REJECTED	NO DECISION	APPROVED	REJECTED
Assessors 91	REJECTED	APPROVED	APPROVED	REJECTED
Assessors 92	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 93	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 94	APPROVED	NO DECISION	APPROVED	APPROVED
Assessors 95	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 96	APPROVED	NO DECISION	APPROVED	APPROVED
Assessors 97	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 98	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 99	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 100	APPROVED	NO DECISION	APPROVED	REJECTED
Assessors 101	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 102	APPROVED	APPROVED	APPROVED	APPROVED
Assessors 103	NO DECISION	NO DECISION	NO DECISION	NO DECISION
Assessors 104	NO DECISION	NO DECISION	NO DECISION	NO DECISION
Assessors 105	REJECTED	REJECTED	REJECTED	REJECTED
Assessors 106	REJECTED	REJECTED	REJECTED	REJECTED

Outline of this talk

Problem Description

Cluster Analysis

Artificial Neural Networks

Support Vector Machines

Computational results

Computational results, 1

Computational results, 2

Computational results, 3

Computational results, 4

Computational results, 5

Conclusions

Computational results, 5

The real 100 of the 106 data have also been used in a second set of experiments with ANN and SVM.

Here only two classes are considered (APPROVED/REJECTED).

The training set consists of 90, 80 or 70 different patterns. The performance of ANN and SVM is tested on the remaining 10, 20 or 30 patterns and comparison is made with the proposal of the Panel Leader. Again a MLP is used with 9, 5 and 2 nodes in the input, hidden and output layers, respectively.

	Accuracy SVM		Accuracy ANN	
	Train	Test	Train	Test
90-10	100.00%	90.00%	98.89%	80.00%
	(90/90)	(9/10)	(89/90)	(8/10)
80-20	100.00%	95.00%	98.75%	85.00%
	(80/80)	(19/20)	(79/80)	(17/20)
70-30	100.00%	96.67%	100.00%	86.67%
	(70/70)	(29/30)	(70/70)	(26/30)

[Outline of this talk](#)

[Problem Description](#)

[Cluster Analysis](#)

[Artificial Neural Networks](#)

[Support Vector Machines](#)

Computational results

Computational results, 1

Computational results, 2

Computational results, 3

Computational results, 4

Computational results, 5

[Conclusions](#)

Conclusions

Outline of this talk

Problem Description

Cluster Analysis

Artificial Neural
Networks

Support Vector
Machines

Computational
results

Conclusions

Conclusions and
future directions

Conclusions and future directions

- ⇒ Cluster Analysis, Artificial Neural Networks and Support Vector Machines represent extremely useful methodologies in Sensory Analysis.
- ⇒ Experimental results show that these methodologies can be used in an effective way in the analyzed context.
- ⇒ In the future, we would like to apply this technique to the evaluation of different kinds of food also in the case of temporal series (for example, to recognize the different characteristics of a certain wine such as “Verdicchio di Matelica” along the years “verticale”).

Outline of this talk

Problem Description

Cluster Analysis

Artificial Neural Networks

Support Vector Machines

Computational results

**Conclusions
Conclusions and
future directions**