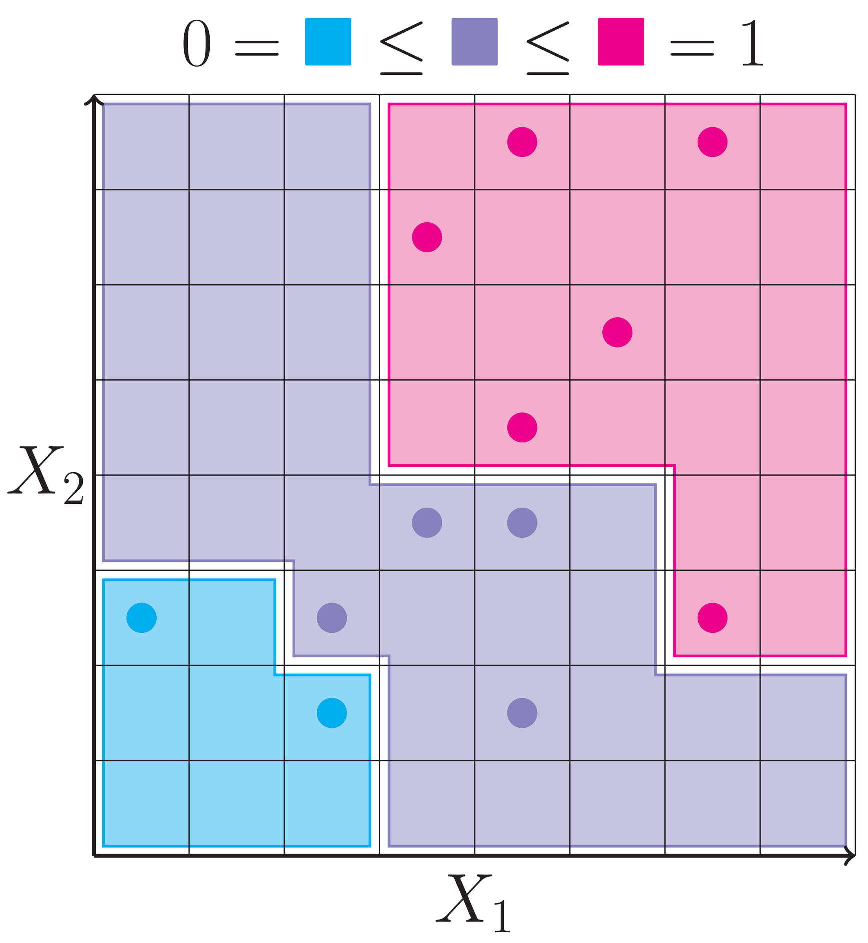
Sugeno Utility Functionals for Monotonic Classication & Decision Rules

Quentin Brabant, Miguel Couceiro

Sugeno Utility Functionals for Monotonic Classification

Quentin Brabant, Miguel Couceiro
quentin.brabant@loria.fr, miguel.couceiro@loria.fr

Monotonic Classification & Decision Rules



Feature space: $\mathbf{X} = X_1 \times \cdots \times X_n$, where X_i is a totally ordered set. Each object is represented by a tuple $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbf{X}$.

Labels: Each object has a label $l(\mathbf{x})$ from a totally ordered set L.

The relation between descriptions and labels is assumed to be **order-preserving**

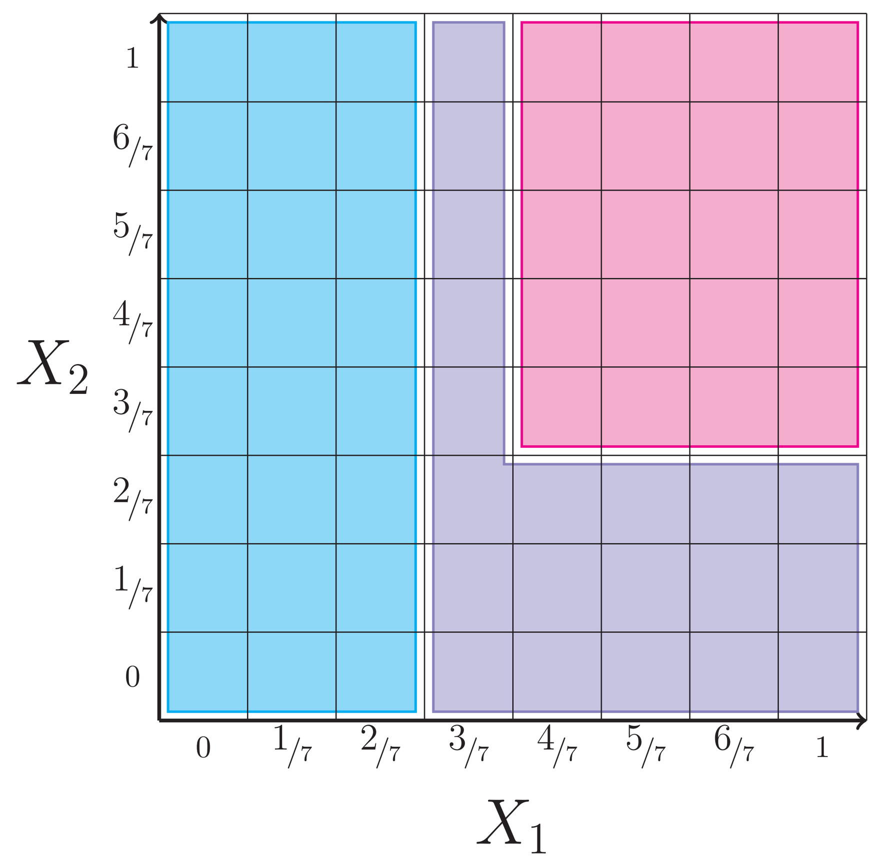
$$a_1 \leq b_1, \ldots, a_n \leq b_n \quad \Rightarrow \quad l(a_1, \ldots, a_n) \leq l(b_1, \ldots, b_n).$$

Aim: to predict the label of objects from their descriptions, with a non-decreasing function $f : \mathbf{X} \to L$.

The function can be specified by decision rules of the form:

$$\forall i \in A, \; x_i \geq \alpha_i \quad \Rightarrow \quad l(\mathbf{x}) \geq \delta,$$

where $A \subseteq \{1, \ldots, n\}$. Sets of such rules can describe any non-decreasing function from $\mathbf{X}$ to L.



Example: These rules express the function at the left.

$x_1 \geq 3/7 \Rightarrow l(\mathbf{x}) \geq$ ■

$x_1 \geq 4/7, \; x_2 \geq 3/7 \Rightarrow l(\mathbf{x}) \geq$ ■

Sugeno Utility Functionals (SUF)

A **capacity** $\mu : 2^{\{1,\ldots,n\}} \to L$ is a set function verifying

- $\mu(\emptyset) = 0$ and $\mu(\{1, \ldots, n\}) = 1$
- $I \subseteq J \Rightarrow \mu(I) \leq \mu(J)$.

The **Sugeno integral** S_μ defined by μ is the **aggregation** function

$$\max_{I \subseteq \{1,\ldots,n\}} \min(\mu(I), \min_{i \in I} x_i).$$

Let $\varphi = (\varphi_1, \ldots, \varphi_n)$, where each mapping $\varphi_i : X_i \to L$ verifies

- $\varphi_i(0) = 0$ and $\varphi_i(1) = 1$
- $a_i \leq b_i \Rightarrow \varphi_i(a_i) \leq \varphi_i(b_i)$.

A SUF is a combination of a Sugeno integral and mappings $\varphi_1, \ldots, \varphi_n$ of the form

$$S_\mu(\varphi_1(x_1), \ldots, \varphi_n(x_n)).$$

A single SUF is less expressive than decision rules. A maximum of several SUFs can represent any set of decision rules.

Application

Maxima of SUFs enable a non-parametric method [1] for monotonic classification.

Principle: To fit the data with a max-SUF using the smallest possible number of SUFs.

The max-SUF can then be translated back into rules.

Result: The method is competitive (in terms of accuracy) with state of the art methods [2] for learning decision rules.

References

Try the method on your data: https://github.com/QGBrabant/SUF4OC

[1] Q. Brabant, M. Couceiro, D. Dubois, H. Prade, and A. Rico. Extracting Decision Rules from Qualitative Data via Sugeno Utility Functionals. In *IPMU*, Communications in Computer and Information Science, 253–265. Springer, 2018.

[2] J. Blaszczynski, R. Slowinski, and M. Szelag. Sequential covering rule induction algorithm for variable consistency rough set approaches. *Information Sciences*, 181(5):987–1002, March 2011.