Preference Ranking Learning: Application to a Human Resources Problem.

Massimo Gurrieri, UMONS - MathRo
Massimo.Gurrieri@umons.ac.be

CONTEXT

Nowadays companies rely more and more on skills management in order to have a clearer insight into the skill strengths and weaknesses of employees. Several researches in the field of human resources demonstrate the linkage between temperament and managerial roles competencies of a person.

LABEL RANKING

Let \( \mathcal{X} \subseteq \mathbb{R}^n \) be an input space and \( \mathcal{L} = \{ \lambda_1; \lambda_2; \ldots; \lambda_k \} \) a set of labels.

The goal is to learn a “label ranker” in the form of a map \( \mathcal{X} \rightarrow \mathcal{S}_2 \), where the output space is the set of all total orders (i.e. permutations) over the label set.

Such a map should predict for a new instance \( x \) a ranking over the set of labels:

\[
\lambda_{\tau_1}^{-1}(1) \succ \lambda_{\tau_2}^{-1}(2) \succ \ldots \succ \lambda_{\tau_k}^{-1}(k)
\]

DOMINANCE BASED ROUGH SETS APPROACH

ROUGH SETS: the approximation of classes is made by means of elementary sets:

\[
I_f(x) = \{ y \in \mathcal{U} : f(x, b) = f(y, b), \forall b \in B \}
\]

Lower approximation (certain rules):

\[
\overline{B}(CL_i) = \{ x \in \mathcal{U} : I_f(x) \subseteq CL_i \}
\]

Upper approximation (possible rules):

\[
\overline{B}(CL_i) = \bigcup_{x \in CL_i} I_f(x)
\]

Boundary (approximate rules):

\[
BN_f(CL_i) = \overline{B}(CL_i) - \underline{B}(CL_i)
\]

DOMINANCE BASED ROUGH SET APPROACH (DRSA): it is an extension which takes into account order between classes and can deal with inconsistency w.r.t. the dominance principle:

- Ordered classes \( CL_1, CL_2, \ldots, CL_l \);
- Upward and downward union of classes:

\[
CL_i^+ = \bigcup_{t \leq i} CL_t, \quad CL_i^- = \bigcup_{t \geq i} CL_t, \quad i \in \{1, 2, \ldots, l\}
\]

- Granules of knowledge are used in the approximation:

\[
D^+_i(x) = \{ y \in \mathcal{U} : x D^+_i y \}, \quad D^-_i(x) = \{ y \in \mathcal{U} : x D^-_i y \}
\]

DATASET REDUCTIONS

In order to use DRSA as a classifier, the dataset had to be transformed.

On the one hand, each ranking on competences has been split into pairwise preferences instances and on the other hand each temperament trait score has been expanded into two criteria (a gain and a cost criterion).

A new component has been also added in order to take into account a preference decision \( d \) (greater, less, than, equal to).

INFERRED RULES

Preliminary rules (on pairs of competences) have been obtained by means of a rule extraction algorithm (DOMLEM). This algorithm is based on a sequential covering procedure and generates a minimal set of rules.

Next, new rules (on triplets of competences) have been then created by merging rules on a logic basis.