Active Learning by the Naive Credal Classifier

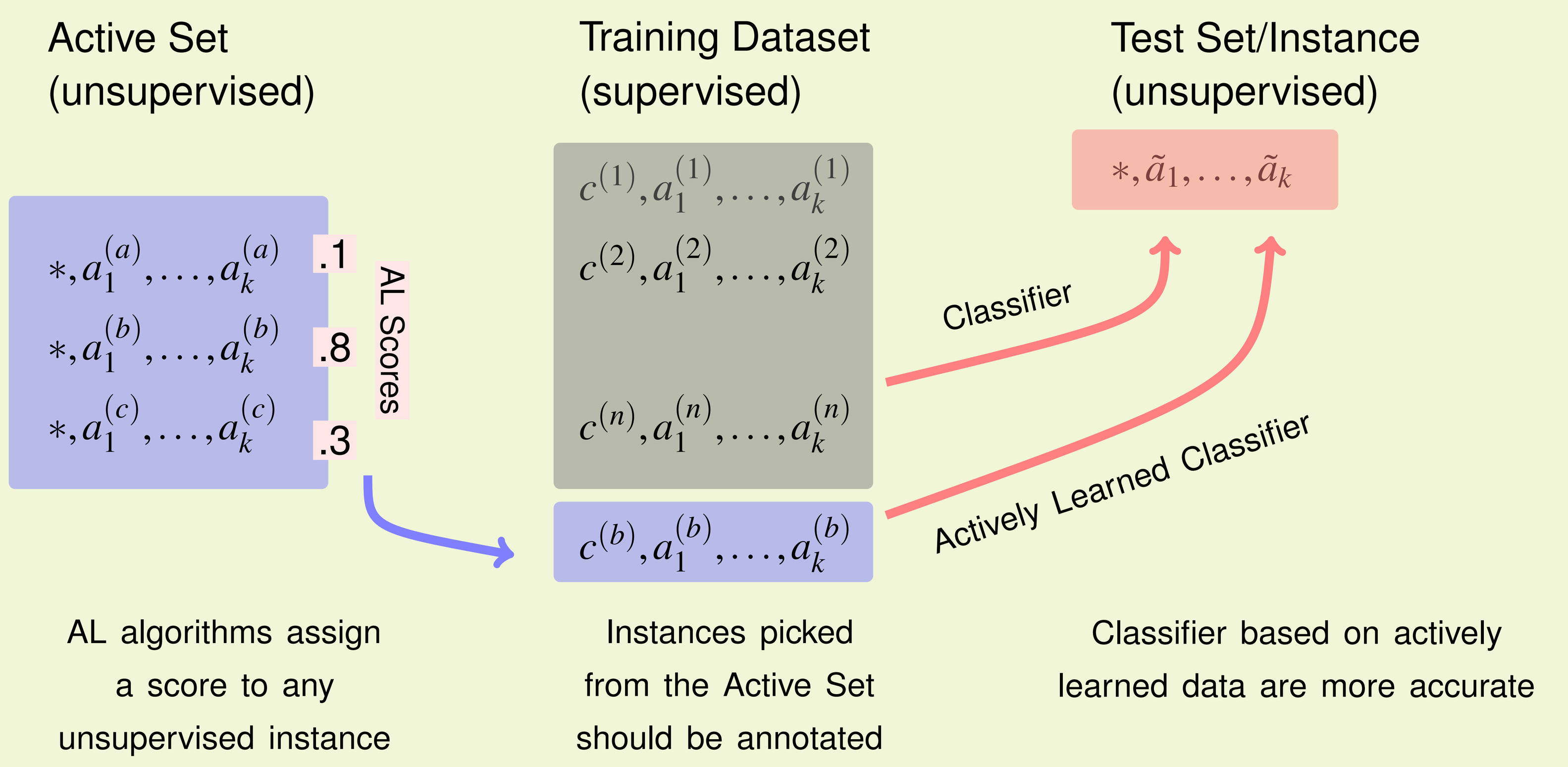
Alessandro Antonucci, Giorgio Corani and Sandra Gabaglio

IDSIA & SUPSI (Switzerland)

{alessandro,giorgio}@idsia.ch sandra.gabaglio@supsi.ch

Active Learning (AL)

Class C taking values in $\mathscr{C}$) Attributes $\mathbf{A} := (A_1, \ldots, A_k)$



Naive Classifiers

Given class C, attributes independent

NAIVE BAYES (NBC)

- A BN quantified from data by a flat Dirichlet prior $\mathrm{Dir}(s\mathbf{t})$
- $P(c,\mathbf{a}) = P(c) \cdot \prod_{i=1}^{m} P(a_i|c)$
- Given test instance $\mathbf{a}$, assigns class $c^* := \arg\max_{c \in \mathscr{C}} P(c|\mathbf{f})$
- $\forall c', c'' \in \mathscr{C}$, dominance test

$$\frac{P(c'|\mathbf{a})}{P(c''|\mathbf{a})} = \frac{P(c',\mathbf{a})}{P(c'',\mathbf{a})} > 1$$

NAIVE CREDAL (NCC)

- BN quantified by set of priors Imprecise Dirichlet model $\mathscr{T} \equiv \{\mathrm{Dir}(s\mathbf{t}) : \mathbf{t} > 0, \sum_i t_i = 1\}$
- A set $\mathscr{C}^* \subseteq \mathscr{C}$ of optimal (undominated) classes
- Conservative dominance test

$$\min_{\mathbf{t} \in \mathscr{T}} \frac{P_{\mathbf{t}}(c',\mathbf{a})}{P_{\mathbf{t}}(c'',\mathbf{a})} > 1$$

Uncertainty Samplings (Bayesian and credal version)

AL score($\mathbf{a}$) should describe how hard-to-classify an instance is
difficult/ambiguous instances give better contribution to the learning

Bayesian Version (US)

- Based on NBC posterior $P(C|\mathbf{a})$
- The smaller the probability of the most probable class, the more hard-to-classify is the instance

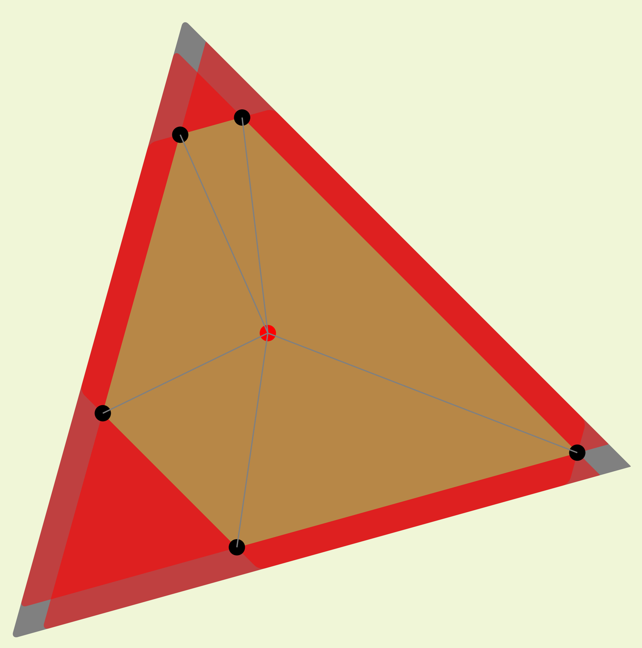
$$\mathrm{score}(\mathbf{a}) \equiv -P(c^*|\mathbf{a})$$

Credal Version (CUS)

- Set of NCC posteriors $\mathscr{P}(C|\mathbf{a})$
- Weaker dominances, instances more hard-to-classify
- Maximum of dominance ratios

$$\mathrm{score}(\mathbf{a}) \equiv -\max_{(c',c'')}\left[\min_{\mathbf{t}} \frac{P(c'|\mathbf{a})}{P(c''|\mathbf{a})}, \min_{\mathbf{t}} \frac{P(c''|\mathbf{a})}{P(c'|\mathbf{a})}\right]$$

Query by Committee (Bayesian and credal version)



Detect posterior(s) robustness

wrt to data resampling (Bayesian)

wrt to different priors (Credal)

Bayesian Version (QbC)

- Training set resampled (by bootstrap) repeat q times
- For each bootstrap, learn NBC from data, evaluate posterior $P_q(C|\mathbf{a})$ (committee member)
- Compute average (center of mass of a polytope in the simplex) $\tilde{P}_{\mathbf{a}}(c) := \frac{1}{q}\sum_{j=1}^{q} P^{(j)}(c|\mathbf{a})$
- Average (KL) distance from CoM $\mathrm{score}(\mathbf{a}) := \frac{1}{q}\sum_{j=1}^{q} \mathrm{KL}[P^{(j)}(C|\mathbf{a}), \tilde{P}_{\mathbf{a}}(C)]$

COMMENT *This is a link between uncertainty measures for sets of distributions and AL (which can be a benchmark for new measures).*

Credal Version (QbC)

- Credal set $\mathscr{P}(C|\mathbf{a})$ of NCC posteriors (committee members)
- Only extreme points matter
- Ignoring committee members in the convex hull? True for QbC! CQbC can only consider the extremes

1. Compute $\{\underline{P}(c|\mathbf{a}), \overline{P}(c|\mathbf{a})\}$ for each $c \in \mathscr{C}$
2. Evaluate the vertices of the consistent credal set ($\underline{P} \leq P \leq \overline{P}$)
3. Apply standard QbC to these vertices

Experiments

constant AL score: random pick among instances in the active set
(variance error decreases, accuracy increases)

AL algorithms should outperform the random pick