Common multilabel formulae

. . . under a common notation

David Charte

Let $A_1, A_2, \dots A_f$ be arbitrary sets. We will call them *input attributes* or simply *attributes*. An instance will take a certain value on each of these sets, that is, we will be working with elements of their cartesian product, $A_1 \times A_2 \times \cdots \times A_f$.

Let L be a finite set. This will be the set of all possible labels. Each instance of a dataset will then have a subset of *active labels*, $Y \subset L$.

Let D be a finite subset of $A_1 \times A_2 \times \cdots \times A_f \times \mathcal{P}(L)$, where $\mathcal{P}(L)$ is the powerset of L, that is, the set of all possible combinations of labels. We will call D a *dataset* and each of its elements an *instance*: $(X, Y) \in D$ where $X \in A_1 \times A_2 \times \cdots \times A_f$ and $Y \in \mathcal{P}(L)$.

Note: Since D is a set, we will be assuming no two instances are identical.