Facultad de Matemática, Astronomía y Física
Universidad Nacional de Córdoba

Inteligencia Artificial 2012

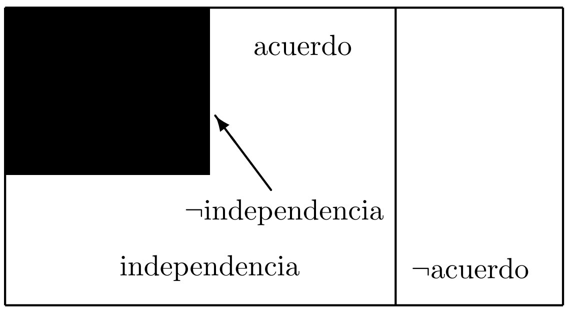
Coeficiente Kappa

El objetivo de esta nota es introducir la definición del Coeficiente Kappa de Cohen.

Dados dos clasificadores que rotulan un conjunto de instancias en categorías mutuamente excluyentes, el *coeficiente kappa* es una medida del porcentaje de *acuerdo* entre los clasificadores bajo la condición de *no independencia*.

Desde un punto de vista probabilístico riguroso, el *coeficiente kappa* es la probabilidad condicional de acuerdo entre los clasificadores, dado que las clasificaciones no son independientes entre sí, es decir que están correlacionadas,

$$\kappa = P(\text{acuerdo} \,|\, \text{no independencia}) = \frac{P(\text{acuerdo} \cap \neg\text{independencia})}{P(\neg\text{independencia})}\,. \tag{1}$$



Teniendo en cuenta que el evento (acuerdo), que contiene las instancias en las que los clasificadores coinciden, puede escribirse como unión de los eventos mutuamente excluyentes (acuerdo $\cap$ independencia) y (acuerdo $\cap\,\neg$independencia) se tiene que

$$P(\text{acuerdo}) = P(\text{acuerdo} \cap \text{independencia}) + P(\text{acuerdo} \cap \neg\text{independencia})\,. \tag{2}$$

Es importante destacar que la condición de independencia es sólo de interés bajo la condición de acuerdo; i.e., (independencia) $\subset$ (acuerdo). Por lo tanto,

$$P(\text{acuerdo} \cap \text{independencia}) = P(\text{independencia})\,.$$

De esta manera, a partir de la Ec. (1) se obtiene

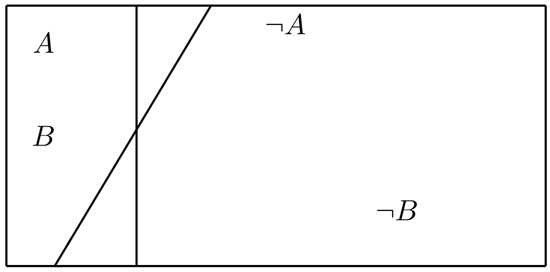
$$\boxed{\kappa = \frac{P(\text{acuerdo}) - P(\text{acuerdo} \cap \text{independencia})}{1 - P(\text{acuerdo} \cap \text{independencia})}} \tag{3}$$

En el caso que los clasificadores sean completamente independientes, es decir,

$$\neg\text{independencia} = \emptyset \quad \text{o bien} \quad \text{acuerdo} = \text{independencia}\,;$$

se tiene que $P(\text{acuerdo} \cap \text{independencia}) = P(\text{acuerdo})$ y resulta $\boxed{\kappa = 0}$. También en el caso trivial (acuerdo) $= \emptyset$ resulta $\kappa = 0$. Por otro lado, si el acuerdo entre los clasificadores es completo, es decir, el evento (acuerdo) se extiende sobre todo el conjunto de instancias; resulta $P(\text{acuerdo}) = 1$ y así $\boxed{\kappa = 1}$. Por último, si los clasificadores están perfectamente correlacionados, (acuerdo) $\cap$ (independencia) $= \emptyset$; es decir, $P(\text{acuerdo} \cap \text{independencia}) = 0$, resulta $\kappa = P(\text{acuerdo})$.

Para fijar ideas, supongamos dos clasificadores A y B, los cuales rotulan las instancias en dos categorías mutuamente excluyentes (positivo y negativo).



El evento con las instancias rotuladas de igual manera por ambos clasificadores puede escribirse como la unión de dos eventos excluyentes

$$\text{acuerdo} = (A \cap B) \cup (\neg A \cap \neg B) = (A \cap B) \cup \neg(A \cup B)\,. \tag{4}$$

De esta manera,

$$P(\text{acuerdo}) = P(A \cap B) + P(\neg A \cap \neg B) = P(A \cap B) + 1 - P(A \cup B)\,, \tag{5}$$

y usando la condición de independencia,

$$P(\text{acuerdo} \cap \text{independencia}) = P(A)\,P(B) + P(\neg A)\,P(\neg B)\,. \tag{6}$$

Usando,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)\,, \tag{7}$$

y

$$P(\neg A)\,P(\neg B) = (1 - P(A))\,(1 - P(B)) = 1 - P(A) - P(B) + P(A)\,P(B) \tag{8}$$

resulta

$$P(\text{acuerdo}) = 1 - P(A) - P(B) + 2\,P(A \cap B)\,, \tag{9}$$

y

$$P(\text{acuerdo} \cap \text{independencia}) = 1 - P(A) - P(B) + 2\,P(A)\,P(B)\,. \tag{10}$$

De esta forma,

$$P(\text{acuerdo}) - P(\text{acuerdo} \cap \text{independencia}) = 2\,(P(A \cap B) - P(A)\,P(B))\,, \tag{11}$$

y así, finalmente se obtiene

$$\boxed{\kappa = 2\,\frac{P(A \cap B) - P(A)\,P(B)}{P(A) + P(B) - 2\,P(A)\,P(B)}\,.} \tag{12}$$

Puede verse de forma directa que si los clasificadores son independientes, $P(A \cap B) = P(A)\,P(B)$, resulta $\kappa = 0$. Mientras que si el acuerdo entre ellos es perfecto, $A = B$, $P(A \cap B) = P(A)$, entonces $\kappa = 1$.