Capturing User reliability

By: Aditya Kumar Akash

1 User reliability

Consider the following matrices based on how a user tags the videos.
For each user u_k, let M^k denote the matrix with rows as video instances and columns are labels. Each element M^k_{ij} denotes would have following values :

$$\begin{aligned} M^k_{ij} &= 1;\ if\ u_k\ tags\ video\ i\ with\ label\ l_j \\ &= -1;\ if\ u_k\ does\ not\ tag\ video\ i\ with\ label\ l_j \\ &= 0;\ \ if\ u_k\ has\ not\ seen\ video\ i \end{aligned} \tag{1}$$

We also have our consensus matrix, M^{cn}, containing the labels for each videos instances.

$$\begin{aligned} M^{cn}_{ij} &= 1;\ if\ consensus\ on\ label\ j\ with\ video\ i \\ &= -1;\ if\ consensus\ on\ label\ j\ not\ with\ video\ i \end{aligned} \tag{2}$$

Now for each user u_k, we find the agreement between the user and consensus model for each label.
The agreement would tell about the reliability of user u_k.

Let us fix label j. For user u_k and label j we construct the confusion matrix using the columns j of M^k and M^{cn}:

$Consensus \downarrow, u_k \rightarrow$	1	-1
1	a	b
-1	c	d

The arrow represents from whom the label is obtained. Thus rows stand for labels obtained from consensus model, while column stand for labels given by user u_i.
Here a stands for number of videos over which user u_k and consensus model has value 1 in the matrix. Similarly we get values for other elements.

Now using *Cohen Kappa*, we get agreement as

$$\kappa = \frac{p_o - p_e}{1 - p_e} = 1 - \frac{1 - p_o}{1 - p_e}, \tag{3}$$

where p_o is the relative observed agreement among raters, and p_e is the hypothetical probability of chance agreement, using the observed data to calculate the probabilities of each observer randomly saying each category. If the raters are in complete agreement then $\kappa = 1$. If there is no agreement among the raters other than what would be expected by chance (as given by p_e), $\kappa \leq 0$.

Let $s = a + b + c + d$. We get $p_o = (a + d)/s$, and $p_e = \frac{(a+b)*(a+c)}{s^2} + \frac{(c+d)*(b+d)}{s^2}$. Putting the values into (3) gives us the agreement, κ^k, for user u_k.
We could call these values as the reliability values of the users for the label j under consideration j.

$$\kappa_j = (\kappa^1, ..., \kappa^k) \tag{4}$$

Overall reliability could be taken as average over the reliability for each label.