Active Testing: Supplementary Materials

Phuc Nguyen [1] **Deva Ramanan** [2] **Charless Fowlkes** [1]

1. Expected Precision and Average Precision

Let Precision and Recall at K be defined as

$$P_k = \frac{1}{k}\sum_{i\leq k} z_i \tag{1}$$

and,

$$R_k = \frac{1}{N_p}\sum_{i\leq k} z_i \tag{2}$$

where N_p is the number of positive instances in the whole set. The average precision is given by integrating precision with respect to recall:

$$\begin{aligned} AP &= \sum_k (R_k - R_{k-1})P_k \\ &= \sum_k \Big(\frac{1}{N_p}z_k\Big)P_k \\ &= \sum_k \Big(\frac{1}{N_p}z_k\Big)\Big(\frac{1}{k}\sum_{i\leq k} z_i\Big) \\ &= \frac{1}{N_p}\sum_k \frac{z_k}{k}\sum_{i\leq k} z_i \end{aligned} \tag{3}$$

We would like to compute expectations when some z_i are unobserved. For notational convenience, let $p_i = P(z_i = 1|\mathcal{O})$ when z_i is unobserved and $\tilde{z}_i$ be the observed value when the ground-truth associated with example i is vetted.

[1]University of California, Irvine [2]Carnegie Mellon University. Correspondence to: Phuc Nguyen <nguyenpx@uci.edu>.

Proceedings of the 35^{th} International Conference on Machine Learning, Stockholm, Sweden, PMLR 80, 2018.

We can then compute expected $Prec@K$ as:

$$\begin{aligned} E[Prec@K] &= \frac{1}{K}\sum_{i\leq K} E[z_i] \\ &= \frac{1}{K}\left(\sum_{i\leq K, i\in V}\tilde{z}_i + \sum_{i\leq K, i\in U} p_i\right) \end{aligned} \tag{4}$$

And the expected change for this metric is given by:

$$\begin{aligned} E_{p(z_i|V)}\left[\Delta_i(z_i)\right] &= p_i\frac{1}{K}|1-p_i| + (1-p_i)\frac{1}{K}|0-p_i| \\ &= \frac{2}{K}p_i(1-p_i) \end{aligned} \tag{5}$$

where we write $p_i = p(z_i = 1|\mathcal{O})$.

Expected AP is more interesting because it includes products of of the z_i.

$$\begin{aligned} E[AP] &= \frac{1}{N_p}\sum_k \frac{1}{k}E[z_k\sum_{i\leq k} z_i] \\ &= \frac{1}{N_p}\sum_k \frac{1}{k}\sum_{i\leq k} E[z_k z_i] \end{aligned}$$

We note that in our application of evaluating instance segmentation, the quantity N_p is known prior to vetting. In other settings, it may also be a random variable that depends on the vetting outcomes. In the following derivation, we temporarily drop the constant $\frac{1}{N_p}$ to reduce notational clutter.

Assuming independence of z_i and z_k, we have:

$$\begin{aligned} E[z_i] &= p_i \\ E[z_i z_k] &= p_i p_k \end{aligned}$$

Expanding the vetted and unvetted terms we can compute:

$$\begin{aligned} E[AP] = &\sum_k \frac{1}{k}\sum_{i\leq k} E[z_k z_i] \\ = &\sum_{k\in V}\frac{1}{k}(\sum_{i\leq k, i\in V} z_k z_i) + \sum_{k\in V}\frac{1}{k}(\sum_{i\leq k, i\in U} z_k p_i) \\ + &\sum_{k\in U}\frac{1}{k}(\sum_{i\leq k, i\in V} p_k z_i) + \sum_{k\in U}\frac{1}{k}(\sum_{i\leq k, i\in U} p_k p_i) \end{aligned}$$

which can be written a bit more compactly as:

$$E[AP] = \Big(\sum_{k \in V} (\frac{z_k}{k} E[Prec@k]) + \sum_{k \in U} (\frac{p_k}{k} E[Prec@k]) \Big)$$

We would like to compute the change in $E[AP]$ when we vet some example. Before vetting sample j, we have:

$$\begin{aligned} E[AP] = \sum_{k \in V} \frac{1}{k} \left[\sum_{i \le k, i \in V} z_k z_i \right] \\ + \sum_{k \in V} \frac{1}{k} \left[\sum_{i \le k, i \in U \setminus j} z_k p_i + z_k p_j \delta[j \le k] \right] \\ + \sum_{k \in U \setminus j} \frac{1}{k} \left[\sum_{i \le k, i \in V} p_k z_i \right] + \frac{1}{j} \left[\sum_{i \le j, i \in V} p_j z_i \right] \\ + \sum_{k \in U \setminus j} \frac{1}{k} \left[\sum_{i \le k, i \in U \setminus j} p_k p_i + p_k p_j \delta[j \le k] \right] \\ + \frac{1}{j} \left[\sum_{i \in U \setminus j, i \le j} p_j p_i + p_j p_j \right] \end{aligned}$$

After vetting the example j, we have:

$$\begin{aligned} E[AP|z_j] = \sum_{k \in V} \frac{1}{k} \sum_{i \le k, i \in V} z_k z_i \\ + \sum_{k \in V} \frac{1}{k} \left[\sum_{i \le k, i \in U \setminus j} z_k p_i + z_k z_j \delta[j \le k] \right] \\ + \sum_{k \in U \setminus j} \frac{1}{k} \left[\sum_{i \le k, i \in V} p_k z_i \right] + \frac{1}{j} \left[\sum_{i \le j, i \in V} z_j z_i \right] \\ + \sum_{k \in U \setminus j} \frac{1}{k} \left[\sum_{i \le k, i \in U \setminus j} p_k p_i + p_k z_j \delta[j \le k] \right] \\ + \frac{1}{j} \left[\sum_{i \le j, i \in U \setminus j} z_j p_i + z_j z_j \right] \end{aligned}$$

The difference between these estimates is,

$$\begin{aligned} \Delta(z_j) &= E[AP|z_j] - E[AP] \\ &= \sum_{k \in V} \frac{1}{k} z_k (z_j - p_j) \delta[j \le k] + \\ &\quad \frac{1}{j} \sum_{i \le j, i \in V} (z_j - p_j) z_i + \\ &\quad \sum_{k \in U \setminus j} \frac{1}{k} \left[p_k (z_j - p_j) \delta[j \le k] \right] + \\ &\quad \frac{1}{j} \left[\sum_{i \le j, i \in U \setminus j} (z_j p_i + z_j z_j - p_j p_i - p_j p_j) \right] \end{aligned}$$

The expected reduction given our estimator for z_j is

$$E[\Delta] \quad = \quad p_j \Delta(z_j = 1) + (1 - p_j) \Delta(z_j = 0)$$

where:

$$\begin{aligned} \Delta(z_j = 0) &= \sum_{k \in V} \frac{1}{k} z_k (-p_j) \delta[j \le k] + \\ &\quad \frac{1}{j} \sum_{i \le j, i \in V} (-p_j) z_i + \\ &\quad \sum_{k \in U \setminus j} \frac{1}{k} p_k (-p_j) \delta[j \le k] + \\ &\quad \frac{1}{j} \left[\sum_{i \le j, i \in U \setminus j} (-p_j p_i - p_j p_j) \right] \end{aligned}$$

$$\begin{aligned} \Delta(z_j = 1) &= \sum_{k \in V} \frac{1}{k} z_k (1 - p_j) \delta[j \le k] + \\ &\quad \frac{1}{j} \sum_{i \le j, i \in V} (1 - p_j) z_i + \\ &\quad \sum_{k \in U \setminus j} \frac{1}{k} p_k (1 - p_j) \delta[j \le k] + \\ &\quad \frac{1}{j} \left[\sum_{i \le j, i \in U \setminus j} (p_i + 1 - p_j p_i - p_j p_j) \right] \end{aligned}$$

Let's just look at the first pair of corresponding terms of $E[\Delta]$,

$$(1-p_j) \sum_{k \in V} \frac{1}{k} z_k (-p_j) \delta[j \le k] + p_j \sum_{k \in V} \frac{1}{k} z_k (1-p_j) \delta[j \le k]$$

It is clear to see that the above summation equals 0. This is also true for the second and third terms,

$$(1-p_j)\frac{1}{j}\sum_{i\leq j, i\in V}(-p_j)z_i + p_j\frac{1}{j}\sum_{i\leq j, i\in V}(1-p_j)z_i$$

$$(1-p_j)\sum_{k\in U\setminus j}\frac{1}{k}p_k(-p_j)\delta[j\leq k]+p_j\sum_{k\in U\setminus j}\frac{1}{k}p_k(1-p_j)\delta[j\leq k]$$

Only the last pair of terms remains, and simplifies as:

$$\begin{aligned} E[\Delta] &= -\frac{1}{j}\sum_{i\leq j, i\in U\setminus j} p_j^3 - p_j + p_j^2(1-p_j) \\ &= -\frac{1}{j}\sum_{i\leq j, i\in U\setminus j} p_j^3 - p_j + p_j^2 - p_j^3 \\ &= -\frac{1}{j}\sum_{i\leq j, i\in U\setminus j} -p_j + p_j^2 \\ &= \frac{1}{j}\sum_{i\leq j, i\in U\setminus j} p_j(1-p_j) \end{aligned}$$

Let r_j be the proportion of unvetted examples scoring higher than example j:

$$\begin{aligned} r_j &= \frac{|\{i\in U : i\leq j\}|}{|\{i\in U\cup V : i\leq j\}|} \\ &= \frac{1}{j}\sum_{i<j}\delta(i\in U) \end{aligned}$$

Putting back in the constant scaling yields the expression given in the paper:

$$\begin{aligned} E[\Delta] &= \frac{1}{N_p}\frac{1}{j}\sum_{i\leq j, i\in U\setminus j} p_j(1-p_j) \\ &= \frac{1}{N_p} r_j p_j(1-p_j) \end{aligned}$$

The term is largest when p_j is 0.5 and decrease as it approaches 0 or 1. The term also decreases when there are many unvetted examples that score higher than j since they have relatively more impact on the AP.

2. Estimator for multilabel classification

Here we derive the basis for Equation 4 in the main paper.

$$\begin{aligned} p(z_i|y_i,s_i) &= \frac{p(z_i,y_i,s_i)}{\sum_{v\in\{0,1\}} p(z_i=v,y_i,s_i)} \\ &= \frac{p(y_i|z_i,s_i)p(z_i|s_i)}{\sum_{v\in\{0,1\}} p(y_i|z_i,s_i)p(z_i|s_i)} \end{aligned}$$

We assume that given the true label, z_i, the observed label y_i is conditionally independent of the classifier score, s_i. With $p(y_i|z_i,s_i) = p(y_i|z_i)$, the expression simplifies to,

$$p(z_i|y_i,s_i) = \frac{p(y_i|z_i)p(z_i|s_i)}{\sum_{v\in\{0,1\}} p(y_i|z_i)p(z_i|s_i)}$$

3. Additional Results

Figure 1 shows the results of estimating absolute precision@48 for the multilabel classification tasks on both NUS-WIDE and Microvideos datasets. In contrast, plots in the main paper show the total absolute error from the true value.

References

Boyd, S., Cortes, C., Mohri, M., and Radovanovic, A. Accuracy at the top. In *Advances in neural information processing systems*, 2012.

Joachims, T. A support vector method for multivariate performance measures. In *Proceedings of the 22nd international conference on Machine learning*, pp. 377–384. ACM, 2005.

Vapnik, V. Principles of risk minimization for learning theory. In *Advances in neural information processing systems*, pp. 831–838, 1992.

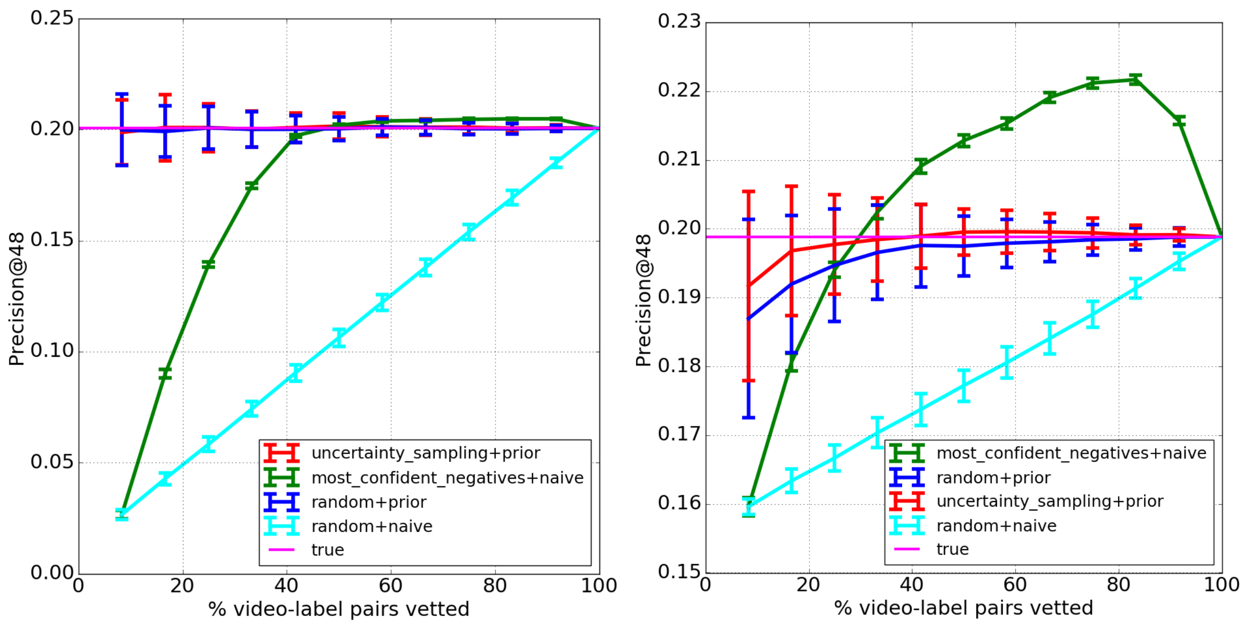

Figure 1. Results for multi-label classification task. The figures show the mean and standard deviation of the estimated Precision@K at different amount of annotation efforts. Using a fairly simple estimator and vetting strategy, the proposed framework can estimate the performance very closely to the true values.