Supplementary Material

A. Proofs

Below are proofs for the regret bounds from Sections 5 and 6.

A.1. Proof of Theorem 2

First, we bound $\mathbb{E}[\| w_{T+1} \|^2]$:

$$\mathbb{E}[w_{T+1}^Tw_{T+1}] = \mathbb{E}[w_{T}^Tw_T + 2w_{T}^T\phi(x_T, \tilde{y}_T) - 2w_{T}^T\phi(x_T, y_T) + \| \phi(x_T, \tilde{y}_T) - \phi(x_T, y_T) \|^2]$$

$$\leq w_{T}^Tw_{T} + 2\sum_{t=1}^{T} \mathbb{E}[w_{t}^T\phi(x_t, \tilde{y}_t) - w_{t}^T\phi(x_t, y_t)] + 4R^2T$$

$$\leq (4R^2 + 2\Delta)T$$

The first line utilizes the update rule from algorithm 2. The second line follows from $\| \phi(x, y) \| \leq R$ and repeating the inequality for $t = T - 1, \cdots, 1$. The last inequality uses the premise on affirmativeness.

Using the update rule again, we get:

$$\mathbb{E}[w_{T+1}^Tw_{*}] = \mathbb{E}[w_{T}^Tw_{*} + (\phi(x_T, \tilde{y}_T) - \phi(x_T, y_T))^T w_{*}]$$

$$= \sum_{t=1}^{T} \mathbb{E}[[U(x_t, \tilde{y}_t) - U(x_t, y_t)]]$$

$$\geq \alpha \sum_{t=1}^{T} \| U(x_t, y^*_t) - U(x_t, y_t) \| - \sum_{t=1}^{T} \xi_t$$

where the last line uses Eq. (4). Using the Cauchy-Schwarz inequality and concavity of $\sqrt{T}$, we get $\mathbb{E}[w_{T+1}^Tw_{*}] \leq \| w_{*} \| \mathbb{E}[\| w_{T+1} \|] \leq \| w_{*} \| \sqrt{\mathbb{E}[\| w_{T+1} \|^2]}$ from which the claimed result follows.

A.2. Proof of Corollary 3

Note that:

$$\hat{y}_t = \text{argmax}_x w_t^T\phi(x, y)$$

Therefore:

$$\forall t, \hat{y}_t : w_t^T\phi(x_t, \hat{y}_t) \leq w_t^T\phi(x_t, \hat{y}_t)$$

Hence:

$$\forall t : \mathbb{E} [w_t^T\phi(x_t, \hat{y}_t)] - \mathbb{E} [w_t^T\phi(x_t, y_t)]$$

$$\leq w_t^T\phi(x_t, \hat{y}_t) - \mathbb{E} [w_t^T\phi(x_t, y_t)]$$

(10)

Given the condition of the corollary, and the above Equation 10, we get that:

$$\frac{1}{T} \sum_{t=1}^{T} \mathbb{E} [w_t^T\phi(x_t, \hat{y}_t)] - \mathbb{E} [w_t^T\phi(x_t, y_t)] \leq \Omega$$

which using Theorem 2 gives us the corresponding regret bound.

A.3. Proof of Theorem 4

This proof is very similar to the one in (Raman et al., 2012), though it solves a different problem. In particular since:

$$\forall t : \mathbb{E} [w_t^T\phi(x_t, \hat{y}_t)] \geq (1 - \beta)w_t^T\phi(x_t, \hat{y}_t)$$

we have that:

$$\mathbb{E}[w_{T}^T(\phi(x_t, \tilde{y}_t) - \phi(x_t, y_t))] \leq \beta w_{T}^T\phi(x_t, \hat{y}_t)$$

From here on, the proof from (Raman et al., 2012) can be used, to prove the corresponding regret bound. Thus in other words, the perturbation can be thought of as a way to produce an $(1 - \beta)-\text{approximate}$ solution to the argmax problem.

A.4. Proof of Proposition 5

Consider the case when documents in positions $i$ and $i + 1$ (call them $d_i$ and $d_{i+1}$) are swapped:\n
$$w_t^T(\gamma_i - \gamma_{i+1})(\phi(x_t, d_i) - \phi(x_t, d_{i+1}))$$

$$\leq \left(1 - \frac{\gamma_i + 1}{\gamma_i}\right)w_t^T(\gamma_i\phi(x_t, d_i) + \gamma_{i+1}\phi(x_t, d_{i+1}))$$

Note that this factor $1 - \frac{\gamma_i + 1}{\gamma_i}$ is largest for $i = 1$. Thus we can state for every swapped pair:

$$w_t^T(\gamma_i - \gamma_{i+1})(\phi(x_t, d_i) - \phi(x_t, d_{i+1}))$$

$$\leq \left(1 - \frac{\gamma_1}{\gamma_1}\right)w_t^T(\gamma_1\phi(x_t, d_i) + \gamma_{i+1}\phi(x_t, d_{i+1}))$$

Summing this over all swapped pairs, and using the fact that each pair has some probability $p$ to be swapped:

$$w_t^T(\phi(x_t, \hat{y}_t) - \mathbb{E}[\phi(x_t, y_t)])$$

$$\leq p \left(1 - \frac{\gamma_1}{\gamma_1}\right)w_t^T\phi(x_t, \hat{y}_t)$$

A.5. Proof of Proposition 6

We prove a more general proposition here:

**Proposition 7** For $\Delta \geq 0$, dynamically setting the swap prob. of 3PR to be

$$p_t \leq \max \left(0, \min \left(1, c(\Delta \cdot t - R_t)\right)\right),$$

(11)

---

2 This holds assuming the inner products with documents are non-negative. Thus algorithmically this can be implemented by only ranking documents with non-negative scores.
Stable Coactive Learning via Perturbation

for some positive constant \( c \), has regret

\[
\frac{1}{\alpha T} \sum_{t=1}^T \xi_t + \frac{4R^2 + 2\Delta + (\gamma_1 - \gamma_2)R\sqrt{4R^2 + 2\Delta}}{T}.
\]

Proof We prove this by using Theorem 2. In particular, we show:

\[
\frac{1}{T} \sum_{t=1}^T w_t^\top (\phi(x_t, \tilde{y}_t) - \phi(x_t, y_t)) < \Delta + \Gamma \sqrt{\frac{4R^2 + 2\Delta}{T}} \tag{12}
\]

where \( \Gamma = (\gamma_1 - \gamma_2)R \). We will show this holds by induction on \( T \). Note that this condition trivially holds for \( T = 0 \) (base case). Now assume it holds for \( T = k - 1 \). We will show it is true for \( T = k \). Consider the cumulative affirmativeness \( R_k = \sum_{i=1}^{k-1} w_i^\top \phi(x_i, \tilde{y}_i) - w_i^\top \phi(x_i, y_i) \). There are 2 cases to consider:

- \( R_k \geq k\Delta \): If this is the case \( p_k = 0 \) i.e., no perturbation is performed for iteration \( k \) and hence \( y_k = \tilde{y}_k = \arg\max_y w_k^\top \phi(x_k, y) \). Therefore \( w_k^\top (\phi(x_k, \tilde{y}_k) - \phi(x_k, y_k)) \leq 0 \); thus \( R_{k+1} \leq R_k \) and hence the induction hypothesis is satisfied.
- \( R_k < k\Delta \): We have \( \|w_k\| \leq \sqrt{k(4R^2 + 2\Delta)} \) as shown in the proof of Thm 2. As per the perturbation, for all \( y_k \) we have \( \|\phi(x_k, \tilde{y}_k) - \phi(x_k, y_k)\| \leq \Gamma^3 \). Next by Cauchy-Schwarz we get \( w_k^\top (\phi(x_k, \tilde{y}_k) - \phi(x_k, y_k)) \leq \|w_k\|\Gamma \). Thus \( R_{k+1} \leq R_k + \Gamma \sqrt{k(4R^2 + 2\Delta)} \); hence satisfying the induction hypothesis.

Thus the induction holds for \( T = k \). Since equation (12) holds for all \( y_t, \tilde{y}_t \), this condition is also satisfied under expectation (over \( y_t, \tilde{y}_t \)). Hence the condition for Theorem 2 is satisfied, thus giving us the bound. Note that the second term on the RHS of Eq. (12) asymptotically disappears.

B. Additional Details of User Study

The ranking function in the ArXiv search engine used 1000 features which can be categorized into the following three groups.

- Features the corresponded to rank as per query similarity with different components of the document (authors, abstract, article etc..). We used different similarity measures. For each of these document-components and similarity measures, we had multiple features of the form \( \text{rank} \leq a \), where \( a \) was a value we varied to create multiple features (we used 2, 5, 10, 15, 25, 30, 50, 100, 200).
- Second-order features the represented pairwise combinations of rank (for the default similarity measure) for 2 different document-components.
- Query-independent features representing the document age and the document category (e.g. AI, NLP, ML, Statistics etc..).

Our baseline, was a hand-coded solution using 35 features considered the most important by us.