LabelCraft: Empowering Short Video Recommendations with Automated Label Crafting

Yimeng Bai[1], Yang Zhang[1], Jing Lu[2], Jianxin Chang[2],
Xiaoxue Zang[2], Yanan Niu[2], Yang Song[2], Fuli Feng[1]

USTC[1] & Kuaishou[2]

email: baiyimeng@mail.ustc.edu.cn

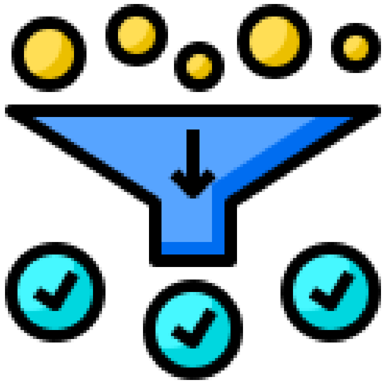
Background

- **Short video recommendation**
  - **Influence**: immense popularity

355.7 million daily active users
612.7 million monthly active users
129.2 minutes daily usage time

- **Method**: personalized video filtering

Video Ocean

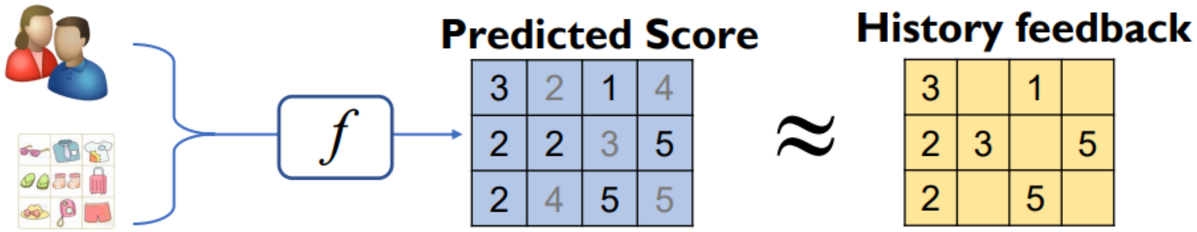


Filtering based on history

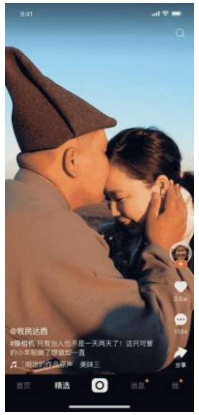
User interested videos

Background

- **Recommendation paradigm**
  - Fitting historical labeled data (**click, purchase**, ….) with a recommender model f



  - Feedback in video recommendation



Implicit feedback

watch time

Explicit feedback

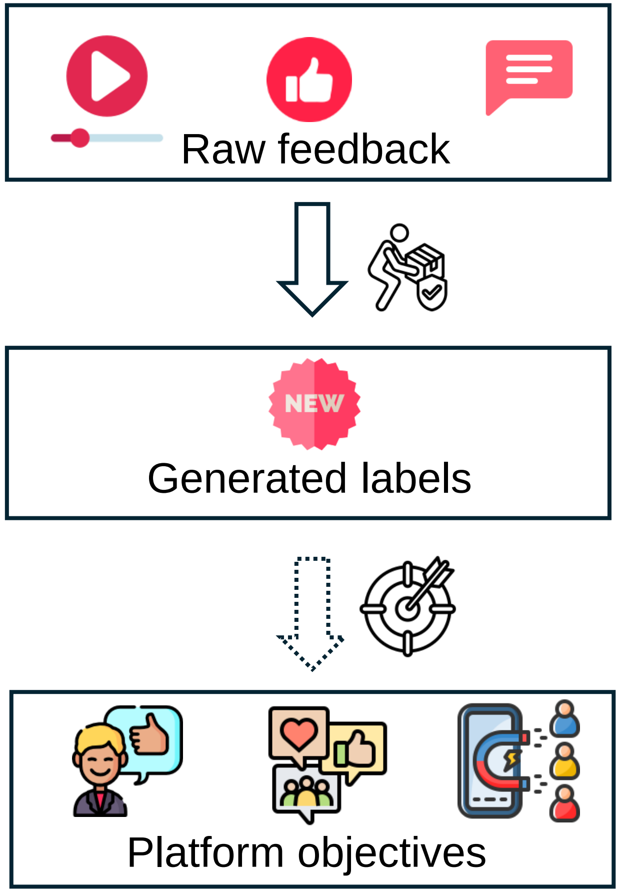
like, follow, forward

Feedback as label directly

- Unreliability
  - 15s/60s v.s. 10s/5s
- Sparsity
  - Explicit but sparse

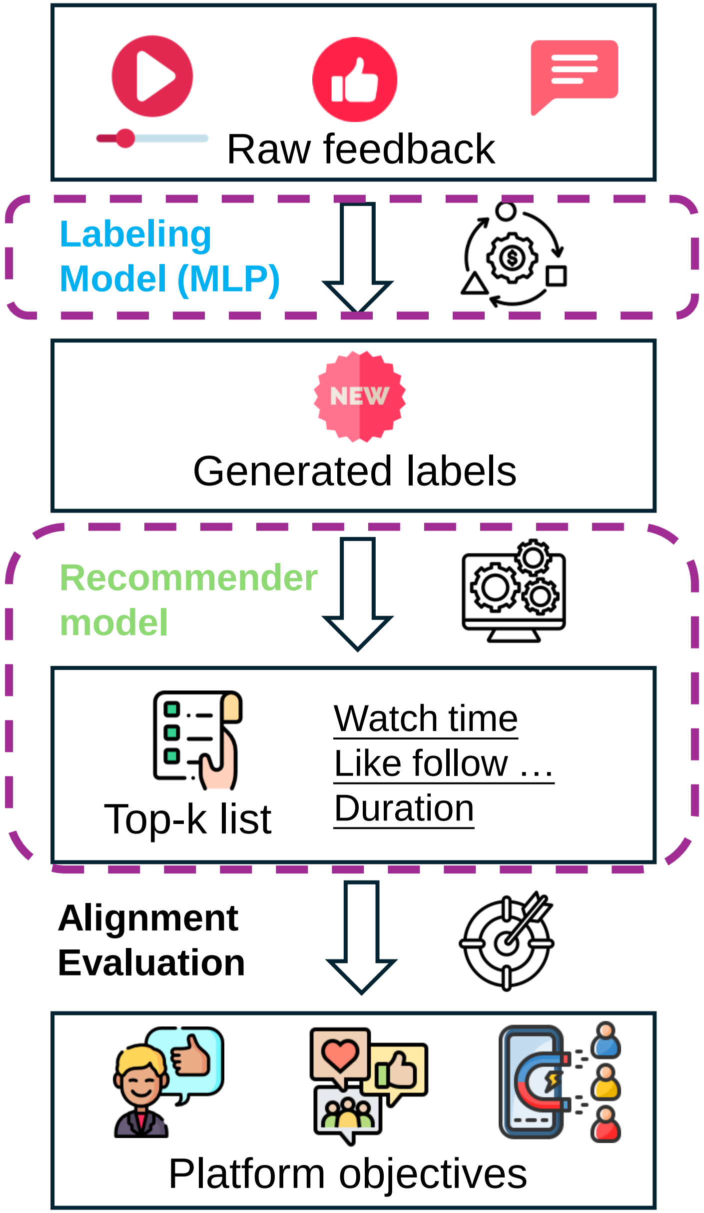
Background

- **Label generation**
  - Map raw feedback to a new label, optimizing platform objectives
  - Examples
    - PlayCompletion = 1 if watch time > duration
    - PlayCompletionRate = watch time / duration
  - Disadvantages
    - Rely on **manual** rules and demand substantial human effort
    - Cannot consider all feedback **comprehensively**
    - **Misalign** with the desired objectives of the platform



Methodology

- **Learnable labeling model**
  - Incorporate both watch time and other feedback
  - Flexible model choice to form complex rules
- **Explicit optimization**
  - Evaluate the label fitted recommender by top-k list
  - Align labeling process with the platform objectives

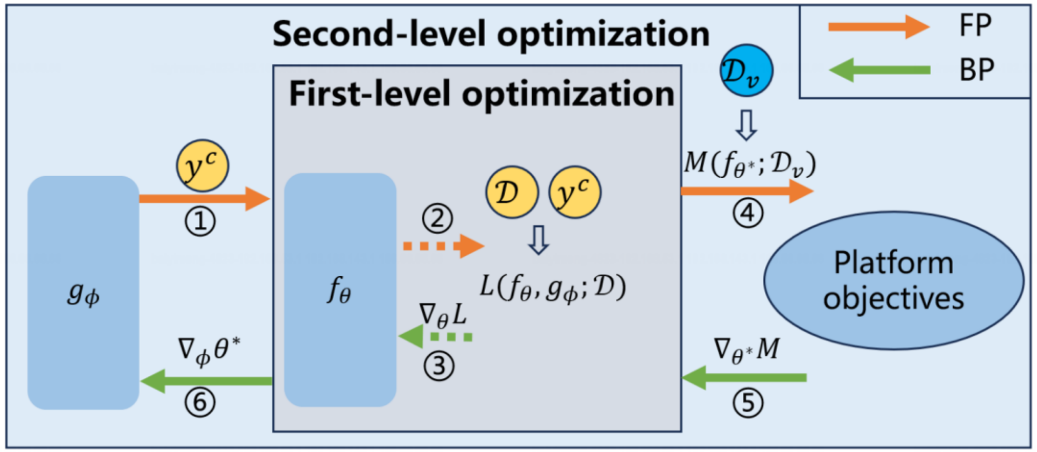


Methodology

- **Problem definition**
  - Labeling model $g_\phi : \mathcal{X} \times \mathcal{Y}^r \longmapsto \mathcal{Y}^c,$ Recommender model $f_\theta : \mathcal{X} \longmapsto \mathcal{Y}^c,$
  - Loss function $L(f_\theta, g_\phi; \mathcal{D}) = \frac{1}{|\mathcal{D}|} \sum_{(\boldsymbol{x}, \boldsymbol{y}^r) \in \mathcal{D}} l(f_\theta(\boldsymbol{x}), y^c) + \lambda \|\theta\|^2, \quad y^c = g_\phi(\boldsymbol{x}, \boldsymbol{y}^r),$
  - Bi-level optimization

$$\boxed{\begin{aligned} &\max_{\phi} \; M(f_{\theta^*}; \mathcal{D}_v) \\ s.t. \quad &\theta^* = \arg\min_{\theta} L(f_\theta, g_\phi; \mathcal{D}), \end{aligned}}$$

Alignment evaluation metric



Methodology

- **Learning strategy**
  - Random initialization of two models
  - Update of **labeling model**
    - Fit the recommender model and obtain a temporary model
    - Evaluate the temporary recommender model on the platform objectives and update the labeling model.
  - Update of **recommender model**
    - Fit new labels and update the recommender model.

Algorithm 1: Training of LabelCraft

Input: Recommender model f_θ, labeling model g_ϕ, training dataset $\mathcal{D}$, hold-out dataset $\mathcal{D}_v$, recommender learning rate η_1 for f_θ, and learning rate η_2 for g_ϕ.

Initialize ϕ and θ randomly;
while *Stop condition is not reached* **do**
  // Step 1 (update of ϕ);
  Compute θ' with $\theta' = \theta - \eta_1 \nabla_\theta L(f_\theta, g_\phi; \mathcal{D})$,
  Update ϕ according to $\phi \leftarrow \phi + \eta_2 \nabla_\phi M(f_{\theta'}; \mathcal{D}_v)$,
  // Step 2 (update of θ);
  Update θ according to Equation $\theta \leftarrow \theta - \eta_1 \nabla_\theta L(f_\theta, g_\phi; \mathcal{D})$,
end
return f_θ, g_ϕ

Methodology

- **Objective representation**
  - Evaluate based on the **top-k list** provided by the recommender model, using **SOFT** top-k [1] technique to ensure differentiability

$$\alpha_{u,\boldsymbol{x}} = SOFT(\boldsymbol{x}; \{f_\theta(\boldsymbol{x}') | (\boldsymbol{x}', \boldsymbol{y}) \in \mathcal{D}_v^u\}) = \begin{cases} 1, & \text{if } f_\theta(\boldsymbol{x}) \text{ is in the top-}k\text{ highest predictions,} \\ 0, & \text{else,} \end{cases}$$

  - Sub-objective
    - **User usage time:** average watch time
    - **User engagement:** average explicit feedback
    - **Duration debias:** std of video duration [2]

$$M_1(f_\theta; \mathcal{D}_v) = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \sum_{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{D}_v^u} \frac{\alpha_{u,\boldsymbol{x}}}{k} scale(y_w),$$

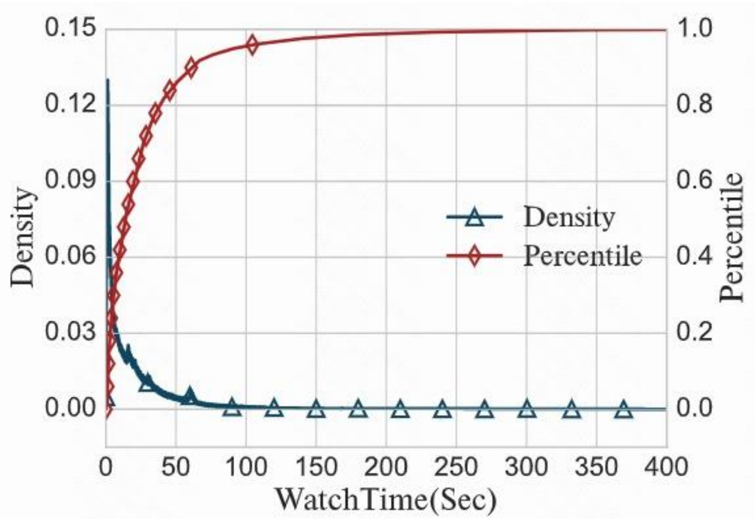
$$M_2(f_\theta; \mathcal{D}_v) = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \sum_{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{D}_v^u} \frac{\alpha_{u,\boldsymbol{x}}}{k} \delta(sum(\boldsymbol{y}_e)),$$

$$M_3(f_\theta; \mathcal{D}_v) = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \left(\sum_{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{D}_v^u} \frac{\alpha_{u,\boldsymbol{x}}}{k} (scale(x_d) - E_{w,k})^2 \right)^{-1/2},$$

[1] Xie et al. Differentiable Top-k with Optimal Transport. NIPS 2020.
[2] Wang et al. Surrogate for Long-Term User Experience in Recommender Systems. KDD 2022

Methodology

- **Objective balancing**
  - **Scaling scheme** [3]
    - M_1, M_2, M_3 are all defined based on feedback, while there exist differences in the magnitude among the feedback
    - We adjust the watch time and video duration to a range of 0-1 based on their distributions
  - **Dynamic balancing**
    - The difficulty of learning varies among different objectives
    - We dynamically allocate weights such that smaller weights are assigned to larger losses



$$scale(y_w) = \begin{cases} \frac{y_w}{w_\beta}\beta', & \text{if } 0 \le y_w \le w_\beta, \\ 1-(1-\beta')\frac{w_{max}-y_w}{w_{max}-w_\beta}, & \text{else,} \end{cases}$$

$$M(f_\theta; \mathcal{D}_v) = \sum_{i=1}^{3} softmax(-\tau M_i(f_\theta; \mathcal{D}_v)) \cdot M_i(f_\theta; \mathcal{D}_v),$$

$$softmax(-\tau M_i(f_\theta; \mathcal{D}_v)) = \frac{\exp(-\tau M_i(f_\theta; \mathcal{D}_v))}{\sum_{j=1}^{3} \exp(-\tau M_j(f_\theta; \mathcal{D}_v))},$$

[3] Sun et al. CREAD: A Classification-Restoration Framework with Error Adaptive Discretization for Watch Time Prediction in Video Recommender Systems. AAAI 2024

Experiment

- **Experiment setting**
  - **Datasets：** Wechat (public) & Kuaishou (private)
  - **Baselines：**
    - WT, EF: feedback as labels directly
    - PC, PCR: manually defined labels
    - D2Q, DVR: label debiasing method
    - WT/D2Q/DVR + EF: multi-task learning method
  - **Evaluation metrics:**
    - Customized NDCG based on watch time (**NWTG@k**) and explicit feedback(**NEG@k**)
    - Standard deviation of video duration (**DS@k**)

Experiment

- **Performance comparison**

Method	Kuaishou						Wechat					
	NWTG@10	RI	DS@10	RI	NEG@10	RI	NWTG@10	RI	DS@10	RI	NEG@10	RI
PC	0.2121	41.6%	15	792.1%	0.7902	3.3%	0.4563	39.0%	10	134.4%	0.7776	7.2%
PCR	0.2493	20.5%	67	105.9%	0.8005	2.0%	0.4125	53.8%	12	100.3%	0.8109	2.8%
WT	0.2939	2.2%	113	21.6%	0.7991	2.2%	0.4972	27.6%	15	58.7%	0.8201	1.7%
D2Q	0.2722	10.4%	122	12.4%	0.7949	2.7%	0.6202	2.3%	23	6.8%	0.8191	1.8%
DVR	0.2814	6.7%	135	1.8%	0.7866	3.8%	0.5300	19.7%	18	32.2%	0.8219	1.4%
EF	0.2557	17.5%	113	21.2%	0.8097	0.8%	0.4593	38.1%	15	65.6%	0.8261	0.9%
WT+EF	0.2631	14.2%	119	15.5%	0.8000	2.1%	0.5195	22.1%	21	17.4%	0.8205	1.6%
D2Q+EF	0.2800	7.3%	111	23.4%	0.7896	3.4%	0.5790	9.5%	22	12.8%	0.8197	1.7%
DVR+EF	0.2876	4.4%	124	10.9%	0.7862	3.9%	0.5698	11.3%	22	11.0%	0.8232	1.3%
LabelCraft	**0.3003**	-	**137**	-	**0.8165**	-	**0.6343**	-	**24**	-	**0.8338**	-

- LabelCraft consistently exhibits superior performance compared to the baselines across all aspects
- This consistent superiority emphasizes the remarkable **alignment** between the labels generated by LabelCraft and the multi-aspect platform objectives

Experiment

- **Ablation study**
  - **Variants**
    - Disable balancing (w/o B)
    - Disable scaling (w/o S)
    - Remove labeling model input (w/o WI DI EI)
    - Remove sub-objective (w/o WO DO EO)
  - **Observations**
    - Disabling the scaling (w/o S) or balancing (w/o B) result in that most metrics decrease
    - Removing any sub-objective (w/o WO DO EO) would lead to decrease in at least one metric, particularly the metric corresponding to the removed sub-objective
    - Removing any input from the labeling model (w/o WI DI EI) would lead to more pronounced decreases on certain evaluation metrics

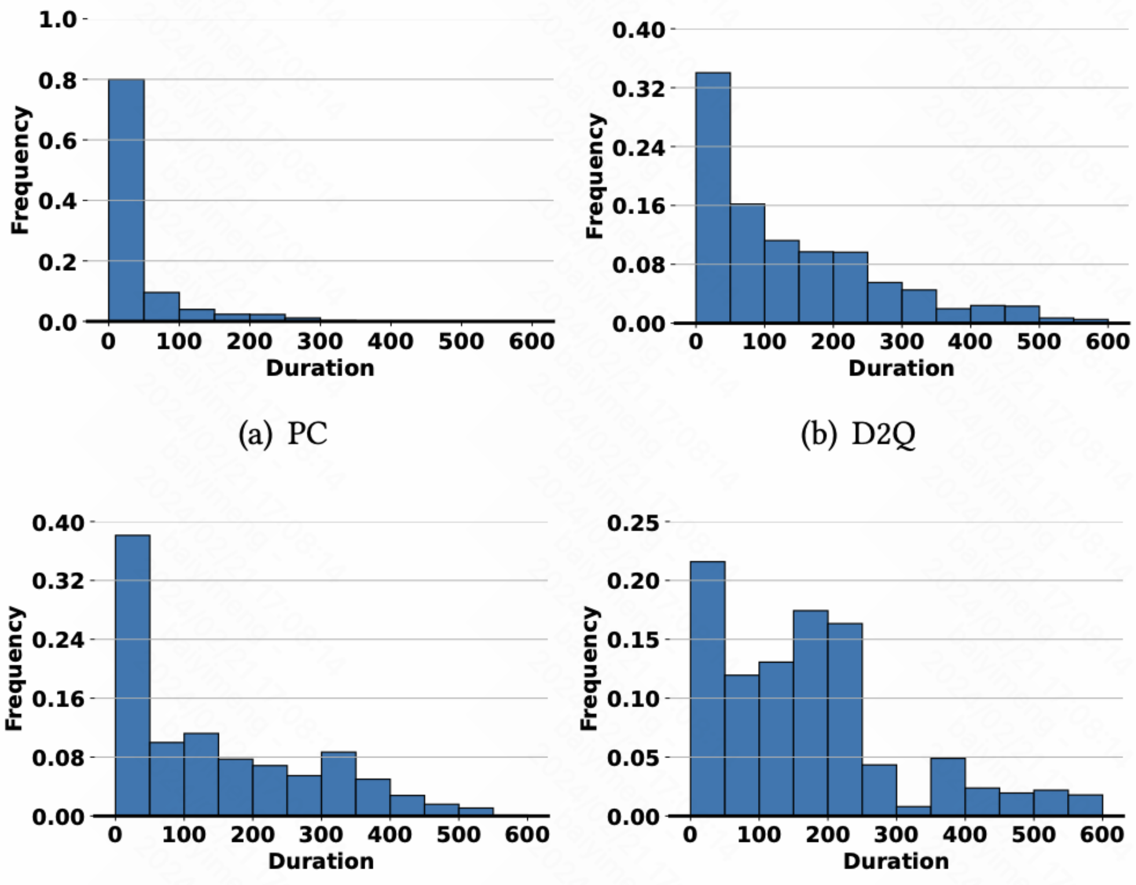
Method	NWTG@10	DS@10	NEG@10
LabelCraft	0.3003	137	0.8165
LabelCraft w/o B	0.2710	122	0.7960
LabelCraft w/o S	0.2627	101	0.8165
LabelCraft w/o WI	$\boxed{0.2677}$	103	0.8151
LabelCraft w/o DI	0.2935	$\boxed{127}$	0.7919
LabelCraft w/o EI	0.3237	124	$\boxed{0.7927}$
LabelCraft w/o WO	$\boxed{0.2785}$	112	0.8001
LabelCraft w/o DO	0.3290	$\boxed{128}$	0.7982
LabelCraft w/o EO	0.3109	131	$\boxed{0.7962}$

Experiment

- **Debiasing performance**

  - LabelCraft **effectively mitigates duration bias**, which can be attributed to the alignment between the labels generated by LabelCraft and the platform objectives

  - When the platform objective is appropriately designed to be free of biases, it could guide the labeling model to generate **bias-free labels**



(a) PC (b) D2Q

(c) DVR (d) LabelCraft

Conclusion & Future Work

- Conclusion
  - Label generation task
  - LabelCraft, automated labeling & explicit optimization

- Future work
  - Align with more complex objectives
  - Better objective balancing method