INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

R&D SHOWCASE 2021

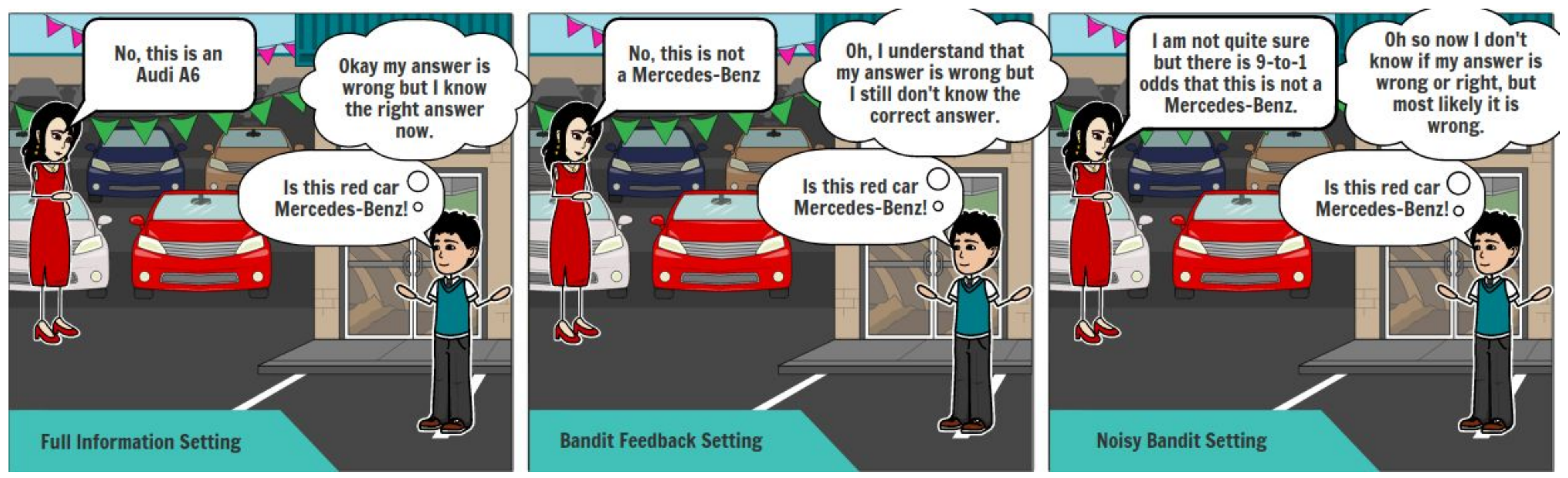
Learning Multiclass Classifier Under Noisy Bandit Feedback

Ad Personalizations

Ad personalization is the act of using customer insights to increase the relevancy of an **ad** to that specific person. These insights could be anything from demographic information to their specific interests, purchasing intent, and buying behavior.

Types of Supervised Learning

- **Full Information Setting:** In this setting, the learner receives the actual class label.
- **Bandit Feedback Setting:** A bandit feedback is revealed to the learner, indicating whether the predicted label is correct or not. This kind of setting can be found in hypothetical recommender systems. Here the user clicks are perfectly correlated to their liking or disliking.
- **Noisy Bandit Setting**: The learner receives noisy bandit feedback (noisy feedback is received by flipping the correct feedback with some small probability). The setting can be found in real-world recommendation systems like Netflix, Google Ads, Youtube etc. This setting also incorporates the noise present in the feedback due to accidental clicking and fake reviews.



Algorithm 3 Noise Rate Estimator (NREst)

Input: $\mathcal{S} = \{(\mathbf{x}^t, \tilde{y}^t), f_\rho^t) \ : \ t = 1 \ldots T\}$

Train a network using $\mathcal{S}$ which approximates $q(\mathbf{x}, \tilde{y}) = \hat{p}(f_\rho = 1|\mathbf{x}, \tilde{y})$

Find $\mathbf{x}^j = \arg\max_{\mathbf{x} \in \mathcal{X}} \ \hat{p}(f_\rho = 1|\mathbf{x}, \tilde{y} = j), \ j \in [K]$

Set $1 - \rho_1 = \hat{p}(f_\rho = 1|\mathbf{x}^l, \tilde{y} = l)$ and $\rho_0 = \hat{p}(f_\rho = 1|\mathbf{x}^k, \tilde{y} = l)$

Output: ρ_0, ρ_1

Google Ads

RCINE Algorithm

Algorithm 2 RCNBF with Implicit Noise Estimation (RCINE)

Input: $\gamma \in (0, 0.5)$, N_s

Initialize: $W^1 = 0 \in \mathbb{R}^{K \times d}, \hat{\rho}_0 = \hat{\rho}_1 = 0, \mathcal{S}$

for $t = 1, 2, \cdots, T$ **do**

Receive $\mathbf{x}^t \in \mathbb{R}^d$.

Set $\hat{y}^t = \arg\max_{r \in [K]}(\mathbf{w}_r^t \cdot \mathbf{x}^t)$

Set $P^t(r) = (1 - \gamma)\mathbb{I}[r = \hat{y}^t] + \frac{\gamma}{K}, \ \forall r$

Randomly sample $\tilde{y}^t$ according to P^t.

Predict $\tilde{y}^t$ and receive feedback f_ρ^t

Calulate $h(f_\rho^t)$ using

$$h(f_\rho^t) = \frac{(1 - \hat{\rho}_{f_\rho^{t'}}) f_\rho^t - \hat{\rho}_{f_\rho^t} f_\rho^{t'}}{1 - \hat{\rho}_0 - \hat{\rho}_1}$$

Define $H^t \in \mathbb{R}^{K \times d}$ such that

$$H_{r,j}^t = x_j^t \left(\frac{h(f_\rho^t)\mathbb{I}[\tilde{y}^t = r]}{P^t(r)} - \mathbb{I}[\hat{y}^t = r] \right)$$

Update: $W^{t+1} = W^t + H^t$

Data: Push $\{(\mathbf{x}^t, \tilde{y}^t), f_\rho^t\}$ in $\mathcal{S}$

if $t\%N_s == 0$ **then**

$\hat{\rho}_0, \hat{\rho}_1 = \text{NREst}(\mathcal{S})$, Clear $\mathcal{S}$

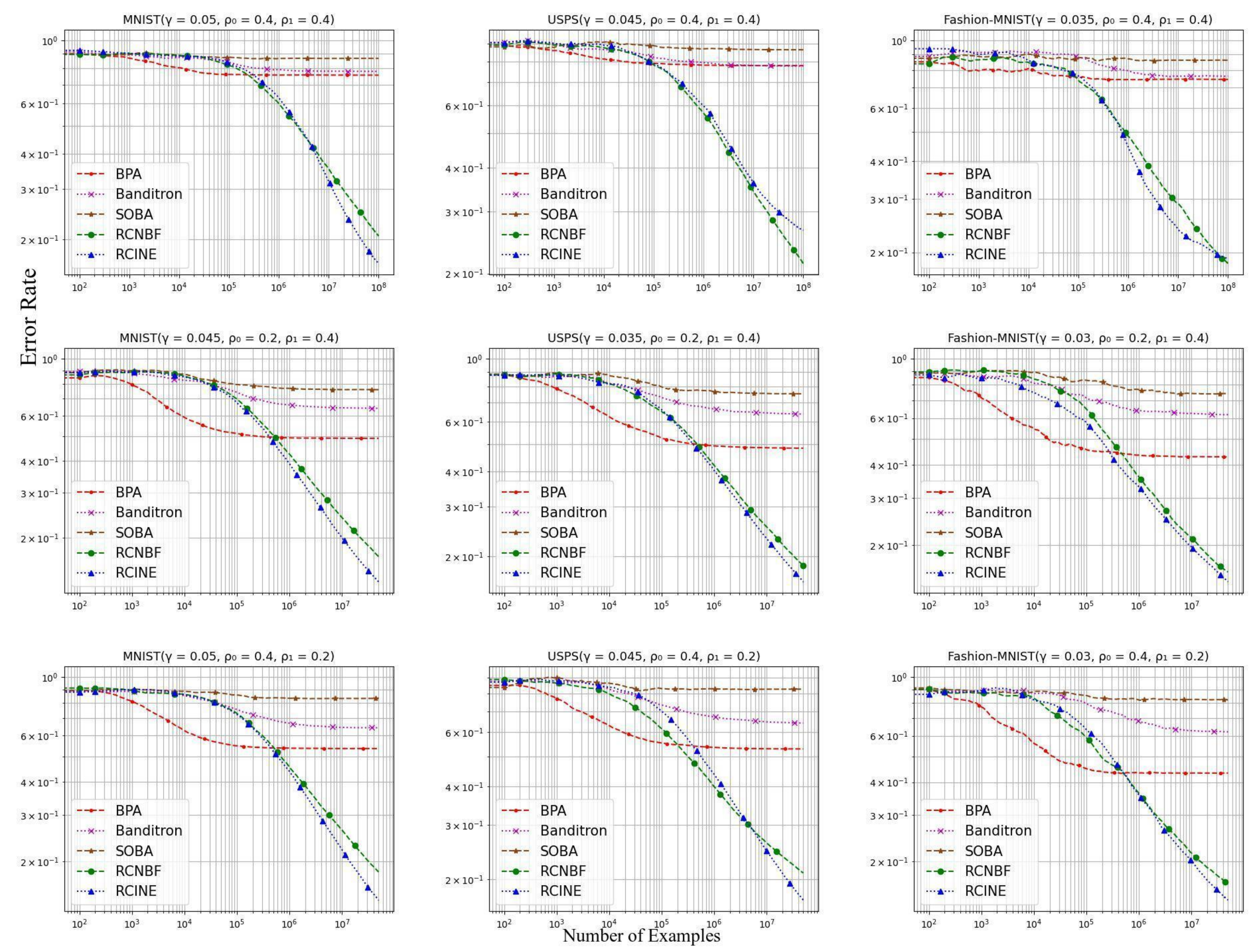
end if

end for

Key Contribution

- The paper also proposes an algorithm for noise rate estimation.
- We also validate our algorithms through experiments on synthetic as well as real-world datasets.
- The paper focuses on making recommendation systems more robust, thus improving recommender systems/ad personalization.
- Some industrial application includes improving content recommendation in OTT platforms (Netflix, Prime) and ad personalization in online advertising platforms (Google Ads)

Results and Simulations



Conclusion

- We propose a robust algorithm for learning multiclass classifiers under noisy bandit feedbacks. The proposed algorithm enjoys mistake bound of the order of $O(\sqrt{T})$ in the high noise case and of the order of $O(T^{2/3})$ in the worst case.

PUBLICATIONS

1. Mudit Agarwal, Naresh Manwani (2021) Learning Multiclass Classifier Under Noisy Bandit Feedback **PAKDD**, 2021.

The code for both the algorithm can be found https://github.com/Mudit-1999/RCINE

Authors: Mudit Agarwal , *Dr.Naresh Manwani*

Research Center: Machine Learning Lab