Appendix for Human-AI Collaboration with Bandit Feedback

Ruijiang Gao[1] , **Maytal Saar-Tsechansky**[1] , **Ligong Han**[2] , **Maria De-Arteaga**[1] ,
Matthew Lease[3] and **Min Kyung Lee**[3]
[1]Dept. of Information, Risk, and Operations Management, University of Texas at Austin
[2]Dept. of Computer Science, Rutgers University
[3]School of Information, University of Texas at Austin
{ruijiang, ml}@utexas.edu, {maytal.saar-tsechansky, dearteaga}@mccombs.utexas.edu,
lh599@cs.rutgers.edu, minkyung.lee@austin.utexas.edu

A Implementation Details of Human Behavior Models in Synthetic Dataset

Black-Box Human Behavior Model To create a black-box human behavior model, we use a random forest classifier on a random 30% subset of the data with full ground-truth labels. After the model is trained, at both training (data generating) and testing time, our synthetic expert will make a random action based on the output probability of the classifier. Since our task is a multi-label classification, the output probability is multiple predictive probabilities of each class which do not sum to 1. In order to make a choice, we use softmax with temperature to normalize the probabilities across classes, the temperatures are set to 10 and 20 for Scene and TMC respectively. When the number of actions increases, the probability output by softmax decreases and a larger temperature can ensure high confidence actions are still selected with relatively high probabilities. Here, high-confidence refers to the original output probability of the black-box model.

Decision Noise Human Behavior Model For decision noise human behavior models, motivated by learning from noisy labels under classification setup [Tanno *et al.*, 2019], we assume each expert will follow a uniform decision accuracy parameterized by ρ. Here we give an example in Table 1. Assume a doctor has $\rho = 0.6$, and needs to diagnose two different patients, Patient 1 and Patient 2, with optimal treatment plan A and B respectively. Table 1 shows the decision probability for this doctor under ρ. In such a setting, the doctor is equally likely to make wrong decisions and have a decision accuracy of 0.6. Similarly, if there are multiple optimal actions, ρ is the summation of the decision accuracy of these actions and they all have the same decision accuracy.

In our experiments, we simulate three expert decision-makers and set ρ to be 0.6, 0.7, 0.8 respectively to see whether our personalization objective can help when experts have diverse skills [Huang *et al.*, 2017].

Optimal Action	A	B	C
Patient 1 - A	0.6	0.2	0.2
Patient 2 - B	0.2	0.6	0.2

Table 1: Decision Accuracy Example

B Dataset Statistics

	# Features	# Labels	Train Size	Test Size
Scene	294	6	1685	722
TMC	30438	22	20017	8579
MLC	1248	6	591	104
Focus	292	2	700	300

Table 2: Dataset statistics.

C Significance Test

We report post-hoc Tukey HSD tests within each cost level for our experiments on the Focus dataset. The significance level is set to 0.05 and the null hypothesis is that the difference between two methods is 0. Results are shown in Table 3. When the human cost is low, the hybrid system demonstrates a significant improvement over the AO baseline. As the human cost increases, the benefit of hybrid team performance over the human-only baseline becomes increasingly significant and the difference between the hybrid team and the AO baseline becomes insignificant.

Cost	H / AO	H / TS	H / JC	H / JCP	AO / TS	AO / JC	AO / JCP	TS / JC	TS / JCP	JC / JCP
0	True	True	False	True	False	True	True	True	True	True
0.05	False	False	False	True	False	True	True	True	True	True
0.1	False	False	True	True	False	False	True	False	True	False
0.3	True	True	True	True	False	False	False	False	False	False
0.5	True	True	True	True	False	False	False	False	False	False

Table 3: Significance Test for Different Reward on Focus

References

[Huang *et al.*, 2017] Sheng-Jun Huang, Jia-Lve Chen, Xin Mu, and Zhi-Hua Zhou. Cost-effective active learning from diverse labelers. In *IJCAI*, 2017.

[Tanno *et al.*, 2019] Ryutaro Tanno, Ardavan Saeedi, Swami Sankaranarayanan, Daniel C Alexander, and Nathan Silberman. Learning from noisy labels by regularized estimation of annotator confusion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 11244–11253, 2019.