Decision Referrals in Human-Automation Teams

Kesav Kaza[1], Jerome Le Ny[1] and Aditya Mahajan[2]
[1]Polytechnique Montreal
[2]McGill University

IEEE CDC 2021
14 December 2021

Motivation

- Automation is coming into the domains of both physical (manufacturing, assembly, etc) and mental tasks (data analysis and decision making)
- Data driven decision support systems (DSS) are an important area of interest in various applications
- Automation and humans have different models of decision making
  - Automation is good at number crunching
  - Humans are good at reasoning with quick mental models
- In collaborative decision making both these strengths can be utilized.

Related work

- Distributed collaborative systems, distributed hypothesis testing with purely automated agents [Tsitsiklis 1993, Tartakovsky et al 2014]
- Dependence of human performance on workload in human factors engineering literature [Tulga and Sheridan 1980, Wickens et al 2015]
- Decision queues where human is modeled as a server with utilization-dependent performance [Jog 2021]
- Task allocation in mixed initiative systems [Hyun et al 2015, Dubois 2020]

Problem setting

- Human-automation team for binary classification tasks
- Hierarchical structure
  - The automation takes first pass at a batch of tasks (say $\mathcal{K}$)
  - It decides which subset $\mathcal{N} \subseteq \mathcal{K}$ of tasks need to be referred to human for review and final decision
  - For all the other tasks ($\mathcal{K} \setminus \mathcal{N}$), the automation makes the final classification decision
- Problem statement : Given a batch (set) of binary classification tasks find the "optimal" subset of tasks to be referred to the human.

Application Scenario

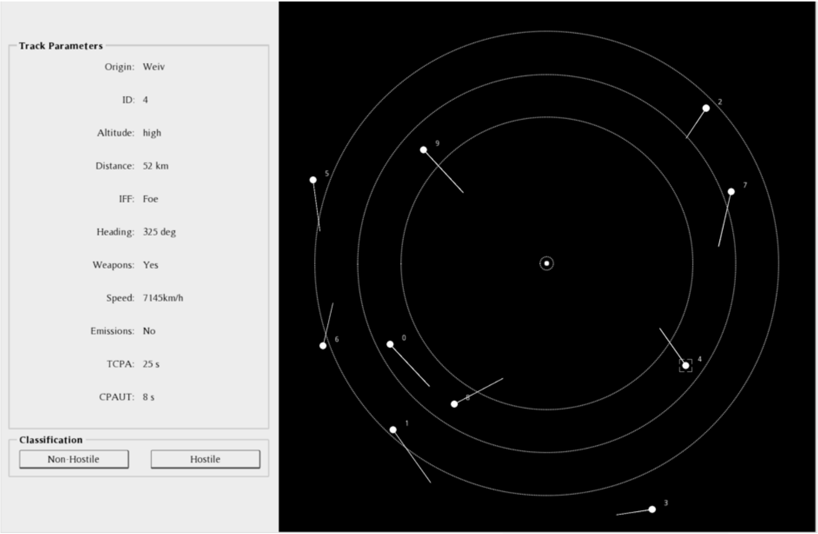

Figure 1: A simulation of a radar screen which shows targets which can be either hostile or non-hostile

System Model

- Batch of K ***i.i.d.*** binary classification tasks, $\mathcal{K} = \{1, 2, \ldots, K\}$.
- Each task $k \in \mathcal{K}$ has true state $H_k \in \{\mathcal{H}_0, \mathcal{H}_1\}$.
- The states are i.i.d. across states with prior $\pi_i = \mathbb{P}(H_k = \mathcal{H}_i), i \in \{0, 1\}$.
- For task k,
  - the automation observes $Y_{1,k} \in \mathcal{Y}_1$
  - the human observes $Y_{2,k} \in \mathcal{Y}_2$
- Observations are random variables which depend on the true state H_k
- Observations are i.i.d. across tasks but conditionally dependent on the states

Observation models

- Automation observation model - Static
  - Conditional distributions over observation values $P_1 : \{\mathcal{H}_0, \mathcal{H}_1\} \to \Delta(\mathcal{Y}_1)$.

$$\mathbb{P}(Y_{1,1}, \dots, Y_{1,K}) = \prod_{k \in \mathcal{K}} \sum_{i \in \{0,1\}} \pi_i P_1(Y_{1,k} | \mathcal{H}_i)$$

  - Example - Let $\mathcal{H}_0 = 0$ and $\mathcal{H}_1 = d_0$. The observations of the automation are given by

$$Y_{1,k} = H_k + N_{1,k}, \quad k \in \mathcal{K},$$

$N_{1,1:K}$ is an independent Gaussian process, independent of $H_{1:K}$, with $N_{1,1:K} \sim Normal(0, \sigma_1^2)$.

Human observation models (Examples)

- Human observation models - Workload dependent
- Workload is defined as the fraction $w = |\mathcal{N}|/|\mathcal{K}| \in [0,1]$ of tasks referred to the human by the automation
- $P_2 : \{\mathcal{H}_0, \mathcal{H}_1\} \times [0,1] \to \Delta(\mathcal{Y}_2)$.

$$\mathbb{P}(\{Y_{2,n}\}_{n\in\mathcal{N}}) = \prod_{n\in\mathcal{N}} \sum_{i\in\{0,1\}} \pi_i P_2(Y_{2,n}|\mathcal{H}_i, w)$$

- *Example 1* - AWGN channel with workload-dependent variance

$$Y_{2,n} = H_n + N_{2,n}, \quad n \in \mathcal{N},$$

- The performance degradation of the human with workload can be captured by assuming that, for some σ_2 such that $\sigma_2^2 \leq \sigma_1^2 < 2\sigma_2^2$,

$$N_{2,n} \sim Normal(0, (1+w)\sigma_2^2), \quad n \in \mathcal{N}.$$

- *Example 2* - AWGN channel with workload-dependent mean

$$Y_{2,n}|\{H_n = \mathcal{H}_0\} \sim Normal(0, \sigma_2^2)$$

$$Y_{2,n}|\{H_n = \mathcal{H}_1\} \sim Normal(d_0(1-w), \sigma_2^2).$$

Human Decision models - Assumptions

- For each task $n \in \mathcal{N}$, the human decides between $\mathcal{H}_0$ and $\mathcal{H}_1$ based only on the observation $Y_{2,n}$
- Human does not have access to the automation's observation $Y_{1,n}$.
- The human also does not account that the automation referred the task after looking at the entire batch

Human Decision models

- The human's classification capability is characterized by the true and false positive rates as function of workload
- When operating at a workload of w, the human's capability is characterized by

$$
\begin{aligned}
P_{2,\mathrm{tp}}(w) &= \mathbb{P}(D_{2,n} = \mathcal{H}_1 | \mathcal{H}_n = \mathcal{H}_1, w), \quad \forall n \in \mathcal{N}, \\
P_{2,\mathrm{fp}}(w) &= \mathbb{P}(D_{2,n} = \mathcal{H}_1 | \mathcal{H}_n = \mathcal{H}_0, w), \quad \forall n \in \mathcal{N}.
\end{aligned}
$$

- Automation does not know the human decision model exactly
- It knows the values of true and false positive rates for each workload level

Problem formulation

- Classification decision costs
  - The cost of final classification decision D_k for task k is

$$\bar{C}(D_k, H_k) = \begin{cases} c_{\text{tp}} & \text{if } (H_k, D_k) = (\mathcal{H}_1, \mathcal{H}_1), \text{true positive} \\ c_{\text{fp}} & \text{if } (H_k, D_k) = (\mathcal{H}_0, \mathcal{H}_1), \text{false positive} \\ c_{\text{tn}} & \text{if } (H_k, D_k) = (\mathcal{H}_0, \mathcal{H}_0), \text{true negative} \\ c_{\text{fn}} & \text{if } (H_k, D_k) = (\mathcal{H}_1, \mathcal{H}_0), \text{false negative.} \end{cases}$$

- Referral decision costs
  - Subset $\mathcal{N} \subseteq \mathcal{K}$ referred to the human
  - The total referral decision cost from the point of view of the automation is

$$|\mathcal{N}| c_m + \sum_{n \in \mathcal{N}} \sum_{i \in \{0,1\}} p^1_{i,n} \bar{C}(D_n, H_i),$$

where $p^1_{i,n}$ is the posterior on the state H_k computed by the automation given the observation $Y_{1,n}$

Optimization problem

- Given the posterior beliefs $\{p^1_{i,k}\}_{k\in\mathcal{K}}, i \in \{0,1\}$ of the automation, and the decision distribution $P_{2,\mathrm{tp}}; P_{2,\mathrm{fp}} : [0,1] \to [0,1]$ of the human, determine $\mathcal{N}$ and $\{D_k\}_{\mathcal{K}\setminus\mathcal{N}}$ so as to minimize the total cost.
- Total cost = Cost of automation classification decisions + Cost of human classification decisions
- Cost of human classification decisions depends on the posterior probabilities of tasks and the true and false positive rates of the human

$$\Gamma_2(\mathcal{N}, w) = \sum_{n\in\mathcal{N}} \Big(p^1_{1,n}[P_{2,\mathrm{tp}}(w)c_{\mathrm{tp}} + (1 - P_{2,\mathrm{tp}}(w))c_{\mathrm{fn}}] + p^1_{0,n}[P_{2,\mathrm{fp}}(w)c_{\mathrm{fp}} + (1 - P_{2,\mathrm{fp}}(w))c_{\mathrm{tn}}] \Big).$$

Optimal Decision Referral Scheme

- G–indices : $G(p_k^1, w) := \bar{C}_1^*(p_k^1) - \bar{\Gamma}_2(p_k^1, w) - c_m.$
- G–index of a task is the cost reduced by referring it to the human

Lemma

For a pre-specified workload $w = |\mathcal{N}|/|\mathcal{K}|$, it is optimal to allocate the tasks with the highest $|\mathcal{N}|$ $G-$indices to the human.

$$\bar{G}(\mathcal{N}) := \sum_{n \in \mathcal{N}} G(p_n^1, |\mathcal{N}|/|\mathcal{K}|). \tag{1}$$

- The total expected cost is equivalent to minimizing

$$\bar{G}^*(w) = \min_{\mathcal{N}:|\mathcal{N}|} = wK, \tag{2}$$

- The optimal workload w can be identified by evaluating $G^*(w)$ for all choices of w.

Numerical Examples

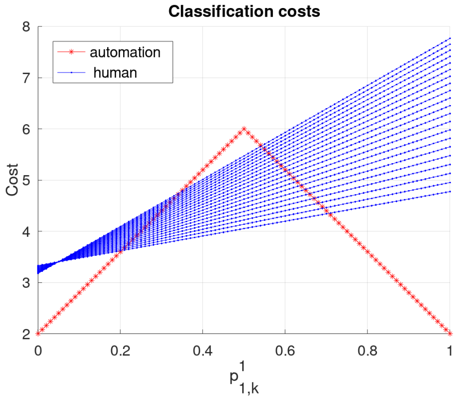

Figure 2: The red hill is the classification cost of the automation, $\bar{C}_1^*(p_k^1)$, as a function of posterior probability $p_{1,k}^1$ of hypothesis $\mathcal{H}_1$. The blue lines show the expected classification cost for the human, $\bar{\Gamma}_2(p_k^1, w)$, $w \in \{1/K, ..., K/K\}$. Batch size $K = 20$. The cost reduction for offloading is $G(p_k^1, w)$, which is the difference between the red and blue functions. ($c_{\mathrm{tp}} = c_{\mathrm{tn}}$ and $c_{\mathrm{fp}} = c_{\mathrm{fn}}$.)

Baseline policies

- **Blind allocation (BA)**, which decides on a workload w^*_{ba} before seeing the batch $Y_{1,1:K}$ and refers $w^*_{\text{ba}}|\mathcal{K}|$ tasks to the human at random.
$$w^*_{\text{ba}} = \arg \min_{w \in \mathcal{W}} \{(1-w)E_1 + wE_2(w)\},$$
- **Static allocation (SA)**, which uses a fixed workload w^*_{sa}, but then refers the tasks in an informed manner according to Lemma 1.

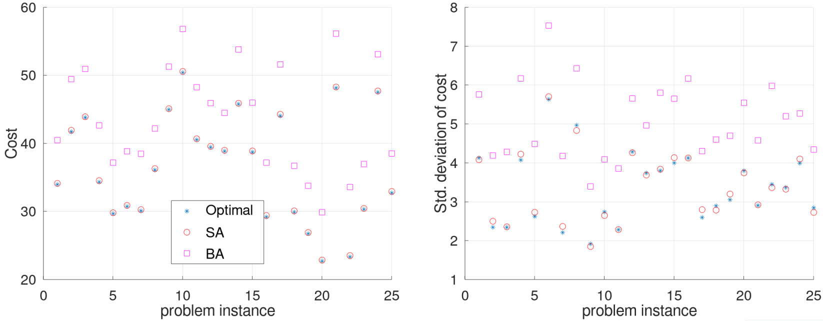

Figure 3: Comparison of various policies for 25 distinct problem instances, for batch size $K = 20$. *[left]* Average cost *[right]* Standard deviations of costs

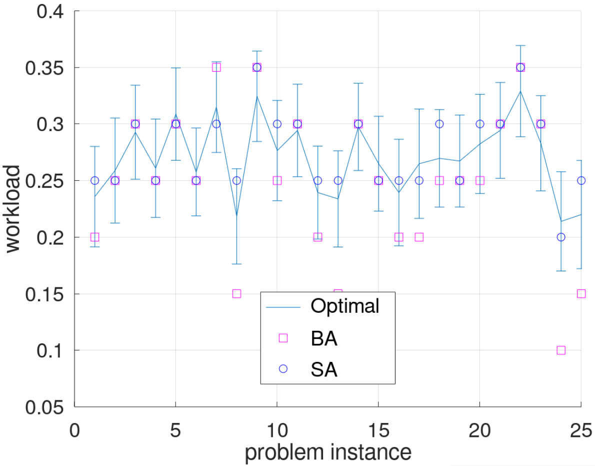

Figure 4: Average workload allotted to human by various policies, over 25 distinct problem instances, for batch size $K = 20$.

Conclusion and Future work

- Informed allocation policies are better than static, blind task allocation schemes
- Informed allocation heuristics which are close to optimal can be devised and employed based on convenience of implementation
- We plan to validate the proposed model through experiments with human participants.
- Other human factors such as fatigue, trust in the automation may be considered.

References

- J. N. Tsitsiklis, “Decentralized detection,” in Advances in Statistical Signal Processing, 1993, pp. 297—344.
- A. Tartakovsky, I. Nikiforov, and M. Basseville, “Sequential analysis: Hypothesis testing and changepoint detection,” CRC Press, 2014.
- M. K. Tulga and T. B. Sheridan, “Dynamic decisions and work load in multitask supervisory control,” IEEE Transactions on Systems, Man, and Cybernetics, vol. 10, no. 5, pp. 217–232, 1980.
- C. D. Wickens, J. G. Hollands, S. Banbury, and R. Parasuraman, “Engineering psychology and human performance,” Psychology Press, 2015.

References

- V. Jog, R. J. La, and N. C. Martins, "Channels, remote estimation and queueing systems with a utilization-dependent component: A unifying survey of recent results," CoRR, vol. abs/1905.04362, 2021.
- B. Hyun, M. Faied, P. Kabamba, and A. Girard, "Mixed-initiative nested classification by optimal thresholding," in 2011 50th IEEE Conference on Decision and Control and European Control Conference, 2011, pp. 7653–7658.
- C. Dubois and J. Le Ny, "Adaptive task allocation in human-machine teams with trust and workload cognitive models," in 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2020, pp. 3241–3246.