Balancing Competing Objectives for Welfare-Aware Machine Learning with Imperfect Data

Esther Rolf†, Max Simchowitz†, Sarah Dean†, Lydia T. Liu†, Daniel Björkegren‡, Moritz Hardt†, Joshua Blumenstock§

†Department of Electrical Engineering and Computer Sciences, University of California, Berkeley ‡Department of Economics, Brown University §School of Information, University of California, Berkeley

Summary

We study algorithmic policies which trade off between two distinct measures of performance. While optimal policies can be described using traditional notions of Pareto optimality when high quality data are readily available, we focus on understanding how to make decisions in noisy or data-poor regimes.

Problem Setting

A central policymaker has two simultaneous objectives: to maximize some private *profit* return (e.g. total user engagement) as well as a public *welfare* objective (e.g. user health):

$$\mathcal{U}_\mathrm{W}(\pi) = \mathbb{E}[w \cdot \pi(x)] \quad \text{and} \quad \mathcal{U}_\mathrm{P}(\pi) = \mathbb{E}[p \cdot \pi(x)]\ .$$

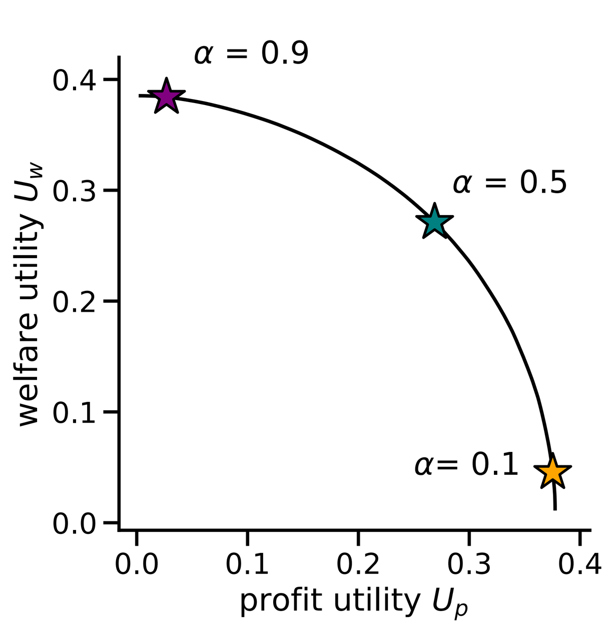
The policymaker makes decisions about *individuals*, who are specified by feature vectors $x \in \mathbb{R}^d$, as well as profit scores $p \in \mathbb{R}$ and change in welfare $w \in \mathbb{R}$, if selected by the policy. Decision policies $\pi(x) \in [0,1]$ corresponding to the *probability* that an individual with features x is selected.

Pareto Optimal Policies

Pareto optimal policies maximize a composite objective:

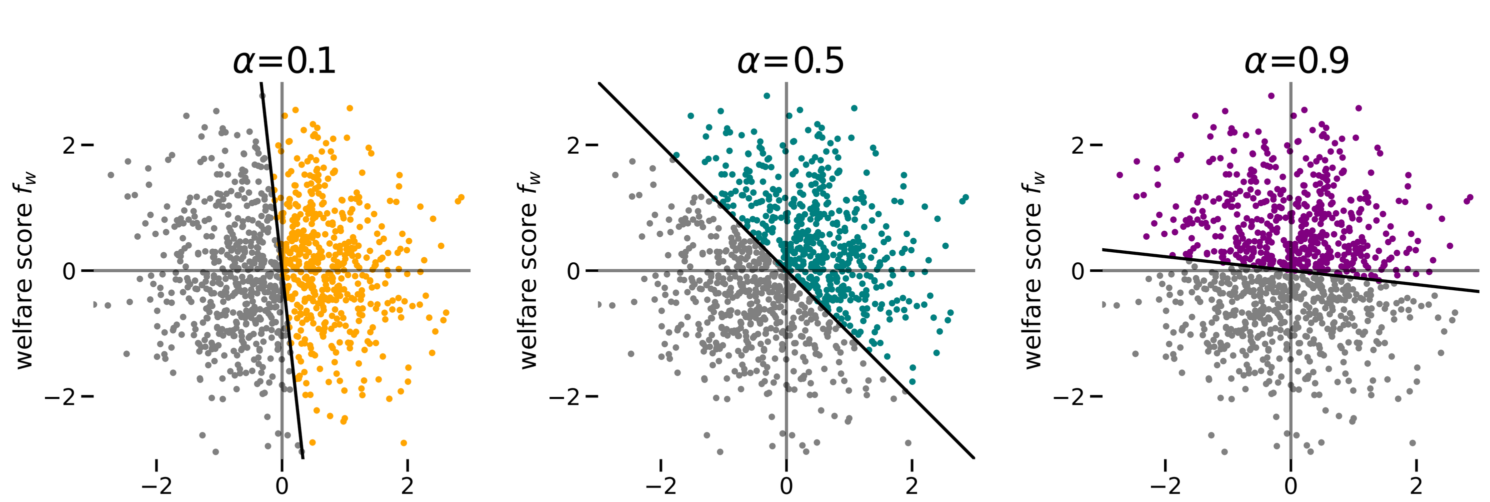
$$\pi^\star_\alpha \in \mathrm{argmax}(1-\alpha)\mathcal{U}_\mathrm{P}(\pi) + \alpha\mathcal{U}_\mathrm{W}(\pi).$$

The parameter α determines this trade-off, tracing the *Pareto frontier*.



When scores are exact, the optimal policy is a threshold:

$$\pi^\star_\alpha(p, w) = \mathbb{I}((1-\alpha)p + \alpha w \geq 0).$$



Optimal Policies from Imperfect Scores

When scores are not exact, but are predictions $\widehat{f}_\mathrm{P}(x)$, $\widehat{f}_\mathrm{W}(x)$ based on features, we consider Pareto optimality with respect to all policies that act on the predicted scores:

$$\Pi_{\mathrm{emp}} = \{\, \pi : (\widehat{f}_\mathrm{P}(X), \widehat{f}_\mathrm{W}(X)) \mapsto [0,1] \,\}\ .$$

The Pareto optimal policy for this class of policies is:

$$\pi^{\mathrm{opt}}_\alpha := \mathbb{I}((1-\alpha)\cdot \overline{\mu}_\mathrm{P} + \alpha \cdot \overline{\mu}_\mathrm{W} \geq 0),$$

a threshold on the conditional expectations:

$$\overline{\mu}_\mathrm{P}(\widehat{f}_\mathrm{P}(x), \widehat{f}_\mathrm{W}(x)) := \mathbb{E}_\mathcal{D}[p \mid \widehat{f}_\mathrm{P}(x), \widehat{f}_\mathrm{W}(x)],$$
$$\overline{\mu}_\mathrm{W}(\widehat{f}_\mathrm{P}(x), \widehat{f}_\mathrm{W}(x)) := \mathbb{E}_\mathcal{D}[w \mid \widehat{f}_\mathrm{P}(x), \widehat{f}_\mathrm{W}(x)]\ .$$

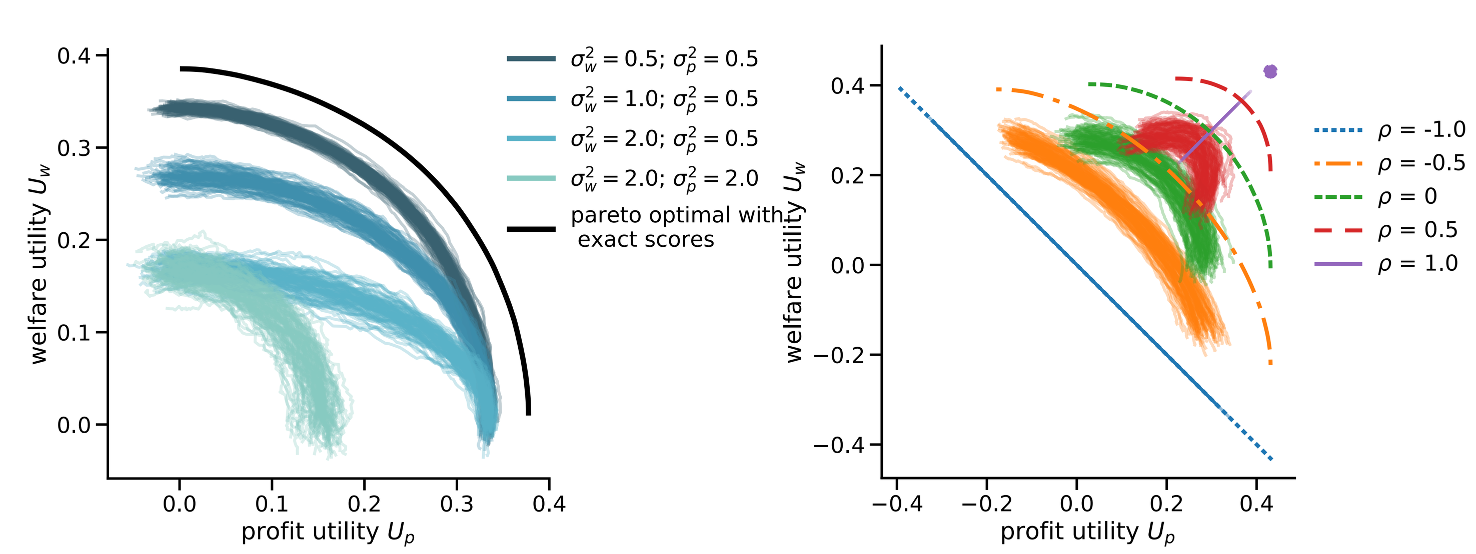
Policies $\pi^{\mathrm{opt}}_\alpha$ for $\alpha \in [0,1]$ trace out an *empirical Pareto frontier*.

- The *empirical* Pareto frontier is dominated by the *exact* Pareto frontier that would arise if exact scores were known.
- Both the empirical and the exact Pareto frontiers exhibit *diminishing marginal returns*: as a policy forgoes more profit to increase welfare, less welfare is gained for the same amount of profit forgone.

Plug-in Policies: When the conditional expectations above are hard to specify, we can define the *plug-in threshold policy*:

$$\pi^{\mathrm{plug}}_\alpha(x) = \mathbb{I}((1-\alpha)\widehat{f}_\mathrm{P}(x) + \alpha\widehat{f}_\mathrm{W}(x) \geq 0)\ .$$

Simulations demonstrate the degradation of the plug-in policy under noise:



(a) Uncorrelated scores with varying additive noise

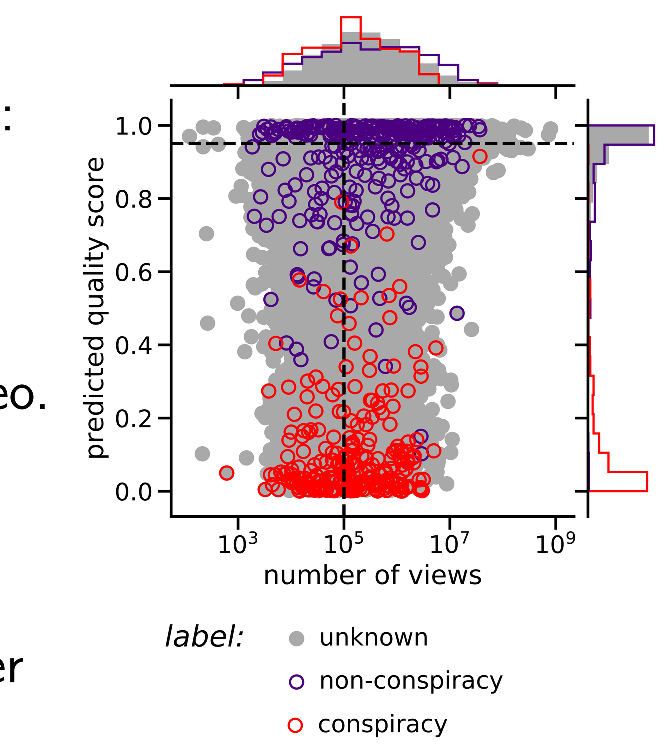
(b) Correlated scores with fixed additive noise

- Pareto optimal policies and plug-in policies coincide when scores are well-calibrated (left above).
- Otherwise, the plug-in policy is not guaranteed to be the optimal policy based on the scores. Under certain settings, we can correct for this using a debiasing procedure on the predicted scores.

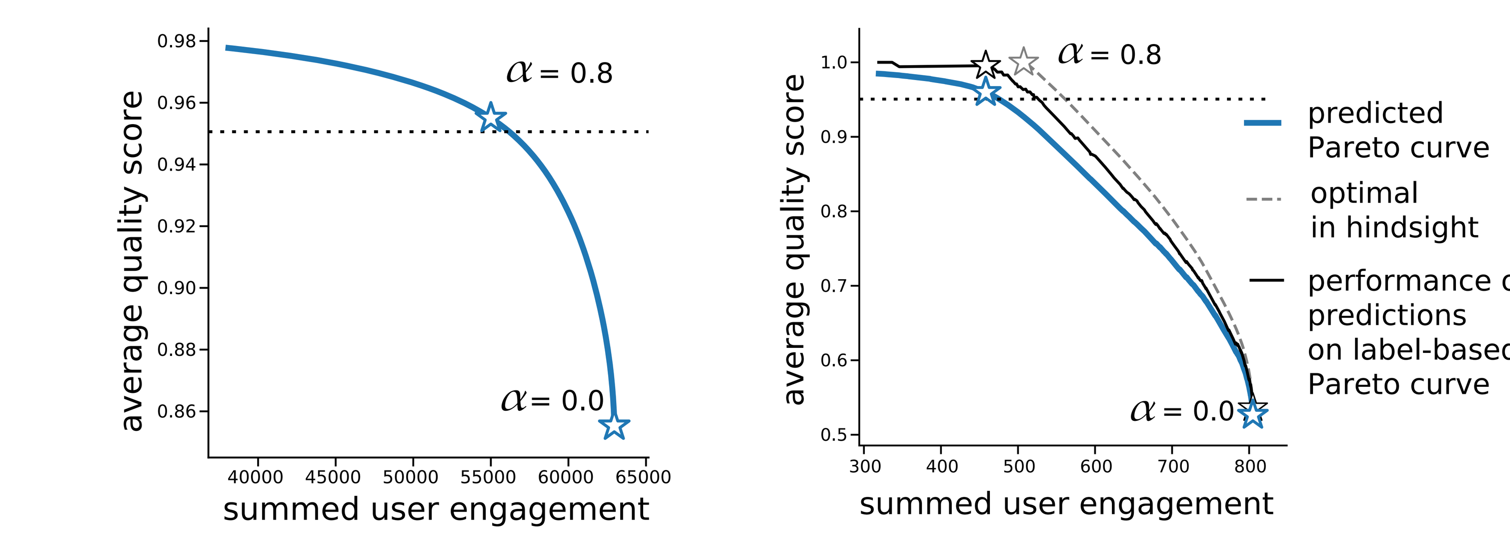
Empirical Results

Youtube Engagement vs Profit:

- 39,817 YouTube videos; policy π decides which videos to show.
- Measure user engagement (p) via number of views of each video.
- Measure quality (w) by predicted scores (empirical) or hand-annotated labels (optimal-in-hindsight) of whether video is a conspiracy video [1].



Using the predicted scores we trace the *predicted* Pareto frontier; using binary ground truth values we trace the *empirical* frontier:



(a) Predicted Pareto Curve

(b) Empirical Pareto Curve

At the maximum-engagement policy, a 1.0% increase in average video quality is achieved with a 0.1% loss in total engagement.

Conclusion

Our framework elucidates trade-offs inherent to optimizing dual objectives with machine learning predictions. Diminishing marginal returns indicate that often a small amount of profit can be sacrificed for large gains in welfare.

References

[1] M. Faddoul, G. Chaslot, and H. Farid. A longitudinal analysis of youtube's promotion of conspiracy videos. *In Preparation*, 2019.

[2] Y. Jin and B. Sendhoff. Pareto-based multiobjective machine learning: An overview and case studies. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 38(3):397–415, 2008.

[3] L. T. Liu, S. Dean, E. Rolf, M. Simchowitz, and M. Hardt. Delayed impact of fair machine learning. In *Proceedings of the 35th International Conference on Machine Learning*, 2018.

Acknowledgments. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE 1752814, the Bill and Melinda Gates Foundation, and the Center for Effective Global Action (Digital Credit Observatory).