Interactive Preference Learning of Utility Functions for Multi-Objective Optimization

Ian Dewancker, Michael McCourt, Samuel Ainsworth

Introduction

In practical machine learning systems, performance tuning is often more nuanced than minimizing a single expected loss, and it may be more realistically discussed as a multi-objective optimization problem

$$\mathbf{f} : \mathcal{X} \to \Omega, \qquad \mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_N(\mathbf{x}) \end{pmatrix}$$

where Ω is the space of possible metric values (in this article, we assume $\Omega \in [0,1]^N$).
We might pose such a problem, involving the accumulation of competing metrics and finding an optimal configuration, as

$$\mathbf{x}_{\text{opt}} = \arg\max_{\mathbf{x}\in\mathcal{X}} \; u(\mathbf{f}(\mathbf{x}))$$

Here, $u : \Omega \to \mathbb{R}$ denotes a *utility function* which encapsulates stakeholders' optimality preferences.

Generative Model for Multi-objective Utility Functions

We propose a model for utility functions with a multiplicative form, more suitable for utility functions with nonlinear structure similar to F-score style utilities designed to balance precision and recall.

$$u(\mathbf{f}(\mathbf{x})) = \prod_{i=1}^{N} u_i(f_i(\mathbf{x}))$$

Each of these individual utility functions could take arbitrary structure; we choose to impose the form of cumulative distribution functions of beta random variables, so that

$$u_i(f_i(\mathbf{x});\, \alpha_i, \beta_i) = \int_0^{f_i(\mathbf{x})} \frac{t^{\alpha_i - 1}(1-t)^{\beta_i - 1}}{B(\alpha_i, \beta_i)}\, \mathrm{d}t$$

$$\log(\alpha_i) \sim \mathcal{N}(\mu_{\alpha_i}, \sigma_{\alpha_i}), \quad \log(\beta_i) \sim \mathcal{N}(\mu_{\beta_i}, \sigma_{\beta_i}) \quad (1)$$

This beta cdf transform has been proposed previously to adapt non-stationary data for use in stationary Gaussian processes [1].

Marginal Likelihood for Binary Preference Data

It may be difficult for stakeholders to judge the utility of a specific set of metrics $\mathbf{f}_1$ in absolute terms, but much simpler to compare a pair $(\mathbf{f}_1, \mathbf{f}_2)$ and decide which of the two is "better" [2]. We allow for two types of observations from users: pairs of multi-objective values where a clear preference in utility is observed (denoted by $\mathcal{D}_P$) and pairs where the utilty is perceived to be equivalent (denoted by $\mathcal{D}_E$),

$$\mathcal{D}_P = \{(\mathbf{f}_1^{p_1} \prec \mathbf{f}_2^{p_1}), \ldots, (\mathbf{f}_1^{p_M} \prec \mathbf{f}_2^{p_M})\}$$
$$\mathcal{D}_E = \{(\mathbf{f}_1^{e_1} \prec\succ \mathbf{f}_2^{e_1}), \ldots, (\mathbf{f}_1^{e_L} \prec\succ \mathbf{f}_2^{e_L})\}$$

We define a parametrization strategy based on marginal likelihood to help find the $4N+1$ best hyperparameters given specific results $\mathcal{D}_P$ and $\mathcal{D}_E$

$$\boldsymbol{\theta} = (\mu_{\alpha_1}, \sigma_{\alpha_1}, \mu_{\beta_1}, \sigma_{\beta_1}, \mu_{\alpha_2}, \cdots, \sigma_E)$$

We also define an auxiliary function

$$u_d(\mathbf{f}_1, \mathbf{f}_2; \boldsymbol{\alpha}, \boldsymbol{\beta}) = \; u(\mathbf{f}_2; \boldsymbol{\alpha}, \boldsymbol{\beta}) - u(\mathbf{f}_1; \boldsymbol{\alpha}, \boldsymbol{\beta})$$

We propose the following two likelihoods:

$$p(\mathcal{D}_P \mid \boldsymbol{\theta}) = \prod_{i=1}^{M} p(u(\mathbf{f}_1^{p_i}) \prec u(\mathbf{f}_2^{p_i}) \mid \boldsymbol{\theta})$$
$$= \prod_{i=1}^{M} \iint h(u_d(\mathbf{f}_1^{p_i}, \mathbf{f}_2^{p_i} \,;\, \boldsymbol{\alpha}, \boldsymbol{\beta}))\, p(\boldsymbol{\alpha}, \boldsymbol{\beta} \mid \boldsymbol{\theta})\, d\boldsymbol{\beta}\, d\boldsymbol{\alpha}$$
$$p(\mathcal{D}_E \mid \boldsymbol{\theta}) = \prod_{j=1}^{L} p(u(\mathbf{f}_1^{e_j}) \prec\succ u(\mathbf{f}_2^{e_j}) \mid \boldsymbol{\theta})$$
$$= \prod_{j=1}^{L} \iint 2\, p(u_E \leq -\left|u_d(\mathbf{f}_1^{e_j}, \mathbf{f}_2^{e_j} \,;\, \boldsymbol{\alpha}, \boldsymbol{\beta})\right|\,)\, p(\boldsymbol{\alpha}, \boldsymbol{\beta} \mid \boldsymbol{\theta}) d\boldsymbol{\beta}\, d\boldsymbol{\alpha}$$

where h is the Heaviside function and $u_E \sim \mathcal{N}(0, \sigma_E)$
We estimate both $p(\mathcal{D}_P \mid \boldsymbol{\theta})$ and $p(\mathcal{D}_E \mid \boldsymbol{\theta})$ using Monte Carlo techniques.

Active Preference Learning of Utility Functions

We adopt a simple active learning algorithm that sequentially decides on a pair of multi-objective values to query the stakeholders for their preference. Our approach draws inspiration from SMBO algorithms, which have been used for human-in-the-loop optimization tasks previously [3].

Algorithm 1 Active Utility Function Pref. Learning

Input: Ω
$\mathcal{D}_P, \mathcal{D}_E \leftarrow$ InitUserPrefs(Ω)
for $i \leftarrow 1$ **to** T **do**
  $\boldsymbol{\theta}_{MLE} \leftarrow \arg\max_{\boldsymbol{\theta}}\, p(\mathcal{D}_P \mid \boldsymbol{\theta})\, p(\mathcal{D}_E \mid \boldsymbol{\theta})$
  $\mathbf{f}_A, \mathbf{f}_B \leftarrow \arg\max_{\mathbf{f}_1, \mathbf{f}_2 \in \Omega}\, a(\mathbf{f}_1, \mathbf{f}_2 \,;\, \boldsymbol{\theta}_{MLE})$
  $p \leftarrow$ GetUserPref($\mathbf{f}_A, \mathbf{f}_B$) ({ A,B,E })
  if $p ==$ E
    $\mathcal{D}_E \leftarrow \mathcal{D}_E \cup (\mathbf{f}_A \prec\succ \mathbf{f}_B)$
  else
    $\mathcal{D}_P \leftarrow \mathcal{D}_P \cup (\mathbf{f}_o \prec \mathbf{f}_p)$
end for

We define an entropy-like acqusition function to guide the search for the pair of metric configurations whose difference in utility has the greatest uncertainty.

$$a(\mathbf{f}_1, \mathbf{f}_2 \,;\, \boldsymbol{\theta}_{\text{MLE}}) = \text{Var}(u_d(\mathbf{f}_1, \mathbf{f}_2 \mid \boldsymbol{\alpha}_{\text{MLE}},\; \boldsymbol{\beta}_{\text{MLE}})),$$

where $\boldsymbol{\alpha}_{\text{MLE}}, \boldsymbol{\beta}_{\text{MLE}}$ follow the distributions defined in (1) with hyperparameters from $\boldsymbol{\theta}_{\text{MLE}}$.

Interactive Tool for Utility Preference Queries

When querying for the stakeholders preferences, a simple back-to-back bar chart of the multi-objective value pairs is displayed as shown in Figure 1.

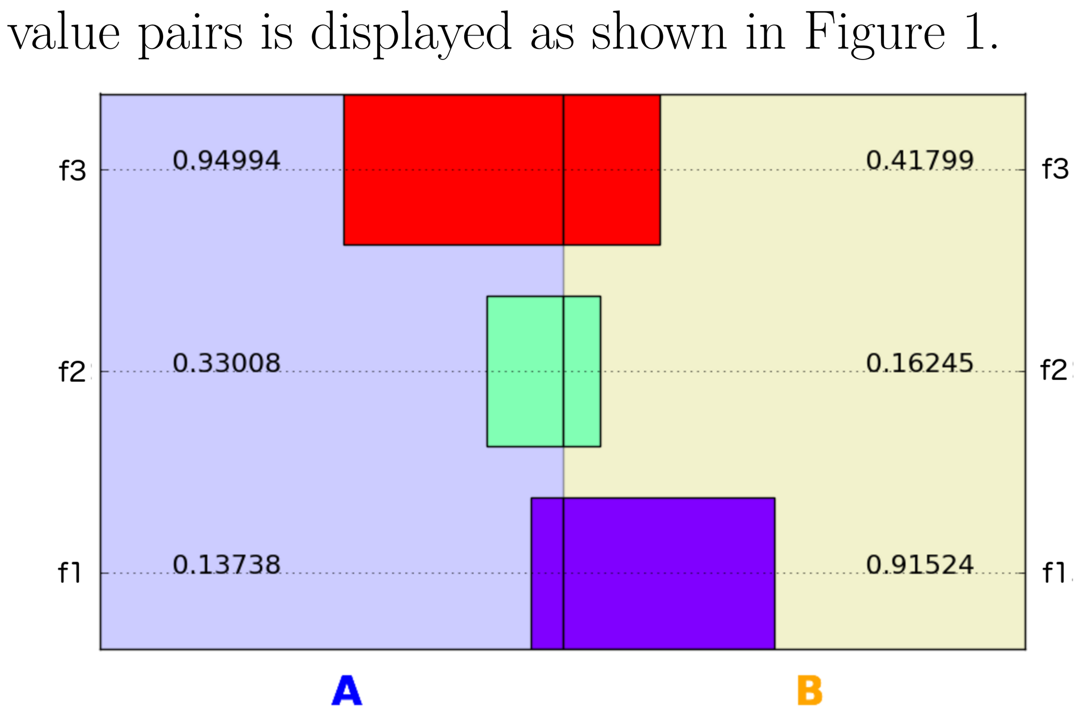

Figure 1: A sample comparison card for preference solicitation showing 3 metrics (f_1, f_2, f_3) in two configurations A and B. Users are asked if they believe utility of configuration A or B to be higher. Users can also specify that they perceive the utilities of the configurations as equal

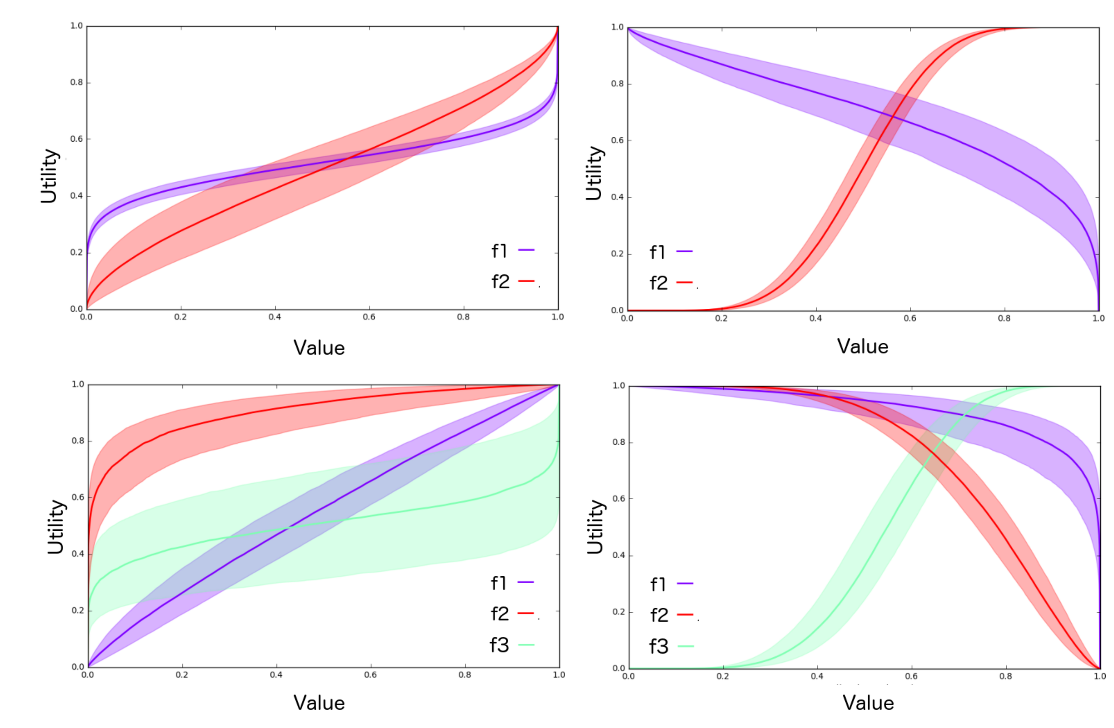

Figure 2: Plots of learned independent utilities with mean and interquartiles. Starting from top left and proceeding clockise, the test utility functions were : 1. max $f_1 + 2f_2$ 2. min f_1 s.t. $f_2 > 0.6$ 3. max $5f_1 + 2f_2 + f_3$ 4. min f_1 s.t. $f_2 < 0.2, f_3 > 0.6$

Since the learned utility is a product of N cdfs, these distributions can be independently visualized as an introspection mechanism.
Figure 2 provides some confidence in the learned joint utility functions. For example, consider the top right plot (2), corresponding to the test utility : min f_1 s.t. $f_2 > 0.6$. We can see that the model has attempted to learn the threshold constraint of $f_2 > 0.6$. We see that the individual utility for f_2 has a sharp, non-linear spike around 0.6. In the full utility function product then, configurations with $f_2 < 0.6$ will be zeroed out and those with $f_2 > 0.6$, the utility function will take on the values of f_1. The individual utility for minimization metrics is defined as the survival function $(1 - u_i(f_i(\mathbf{x}))$ of the beta cdf.

Experimental Results

Explicit test utility functions were used to simulate implicit human utility functions. A hold-out set of 10,000 random multi-objective configurations were generated for each test function and the Kendall rank correlation coefficient was used to quantify the ordinal association between the test utility function values and the mean learned utility function values

Table 1: Kendall-Tau Correlation using Different Search Policies

Test Utility Function	Rnd Search	Pair Entropy
max $f_1 + 2f_2$	**0.8756**	0.8618
max $f_1 + 10f_2$	0.9422	**0.9615**
min f_1 s.t. $f_2 > 0.6$	0.6507	**0.6893**
max $2\, f_1 f_2 \,/\, (f_1 + f_2)$	0.8844	**0.9039**
max $5\, f_1 f_2 \,/\, (4f_1 + f_2)$	0.8949	**0.9120**
max $f_1 + 2f_2 + f_3$	**0.8490**	0.7805
max $5f_1 + 2f_2 + f_3$	**0.8738**	0.8311
min f_1 s.t. $f_2 > 0.6, f_3 < 0.2$	0.2949	**0.3257**
max $2\,(f_1 f_2) \,/\, (f_1 + f_2)$ s.t. $f_3 > 0.95$	0.2309	**0.2648**

References

[1] Jasper Snoek, Kevin Swersky, Richard S Zemel, and Ryan P Adams. Input warping for bayesian optimization of non-stationary functions. In *ICML*, pages 1674–1682, 2014.

[2] Wei Chu and Zoubin Ghahramani. Preference learning with gaussian processes. In *Proceedings of the 22nd international conference on Machine learning*, pages 137–144. ACM, 2005.

[3] Eric Brochu, Tyson Brochu, and Nando de Freitas. A bayesian interactive optimization approach to procedural animation design. In *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 103–112. Eurographics Association, 2010.