# Your Mouse 🖱️ and Eyes 👀 might Determine if the AI Race 👟 will have only One Winner 🥇

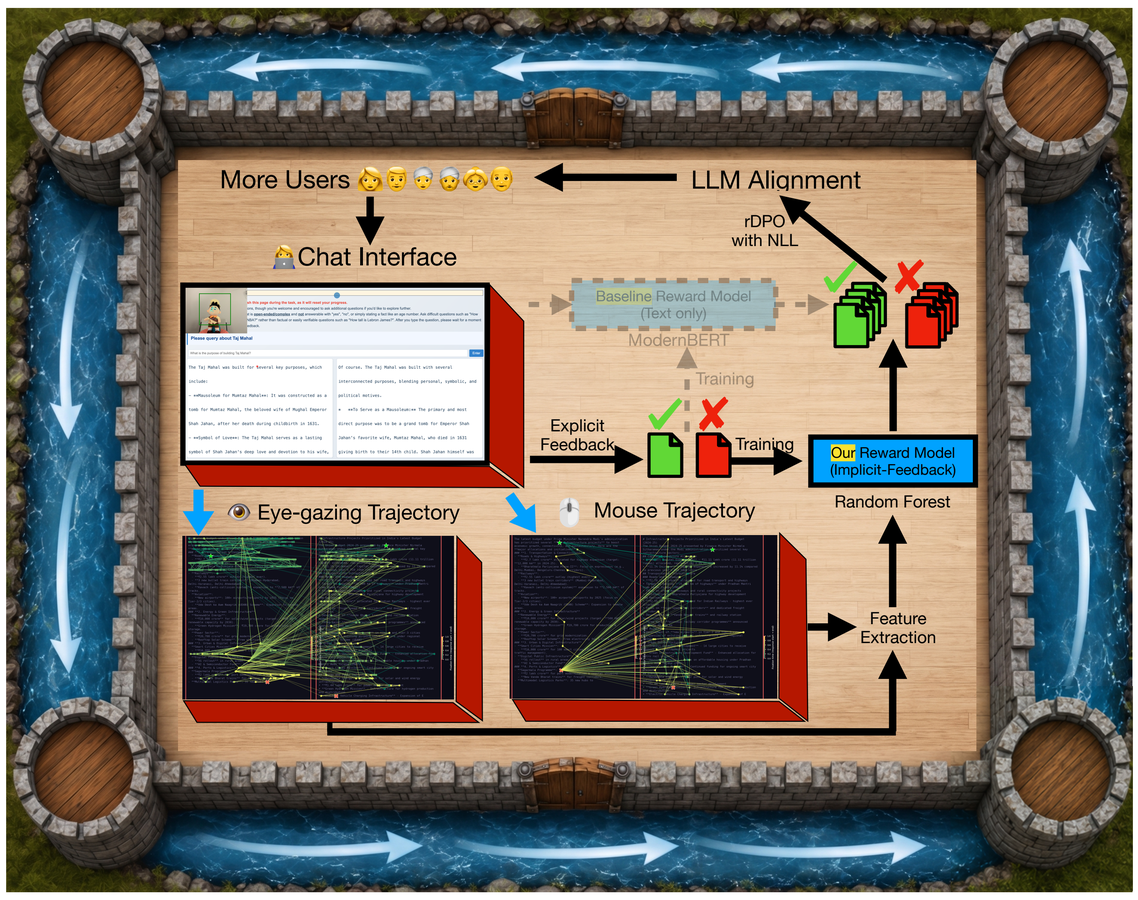
## Your Mouse and Eyes Secretly Leak Your Preference: LLM Alignment using Implicit Feedback from Users

College of Information Science

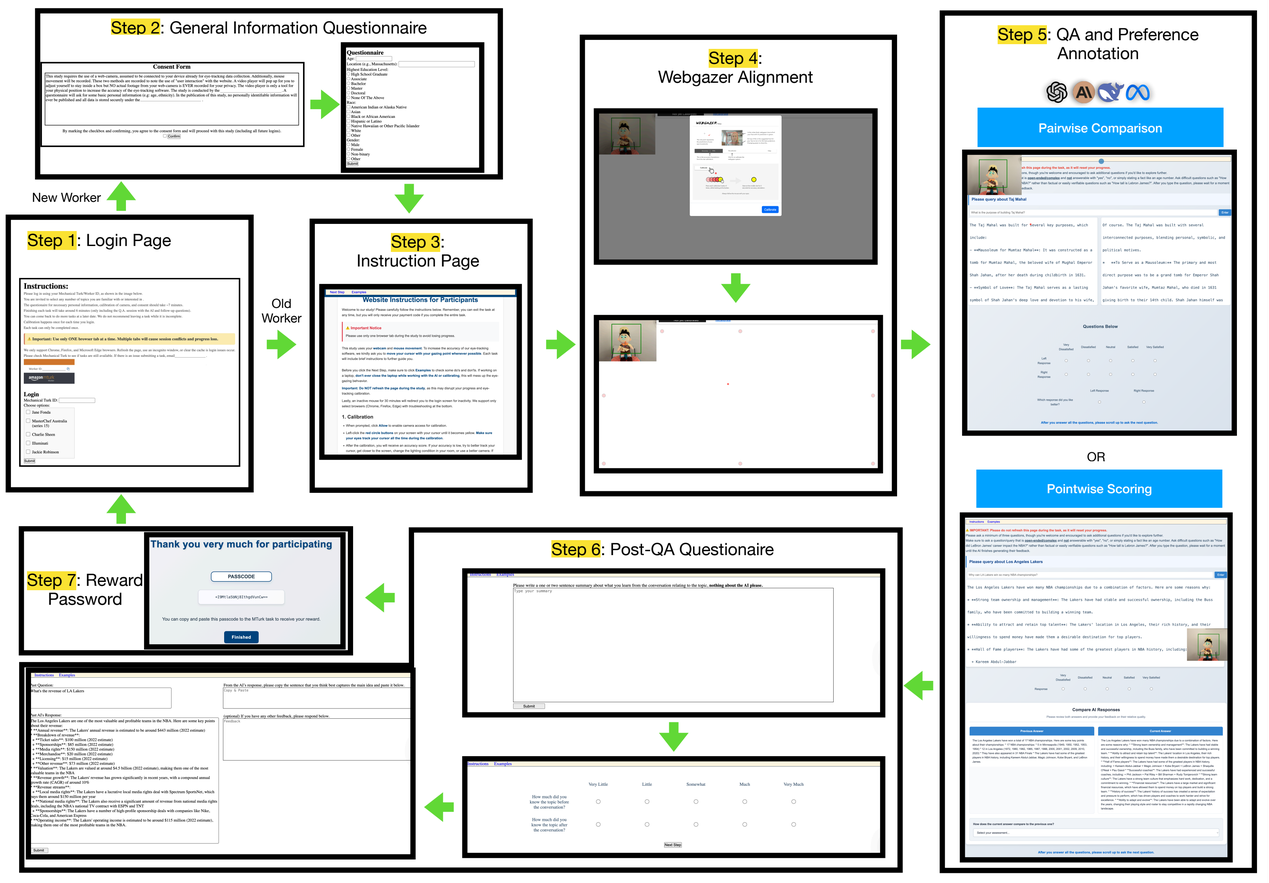
Haw-Shiuan Chang*, Jeffrey Gomez*, Mehul Patwari*, Aryan Sajith, Hamed Zamani

UMASS AMHERST

# Introduction




# Dataset Collection




# Experiment Results on IFLLM (Implicit Feedback for Large Language Models) Benchmark

1336 multi-turn questions from 59 Mechanical Turk workers

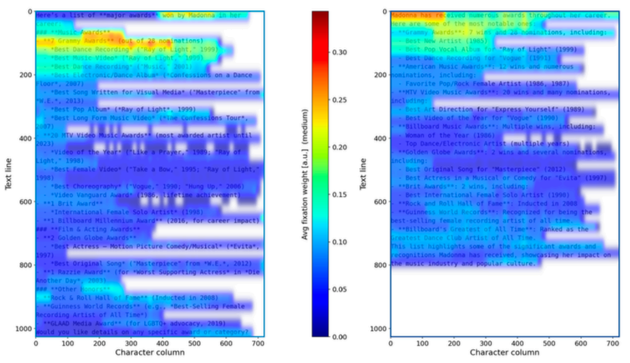


**Figure 3: Average fixation weight over the response text in the pairwise setting, aggregated across all medium-length responses. The displayed text is a randomly selected example.**

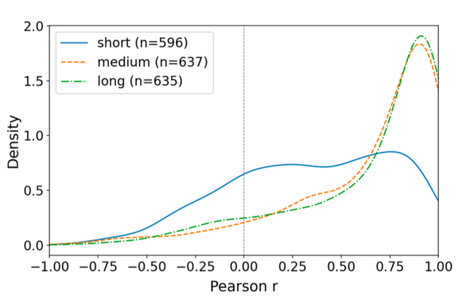
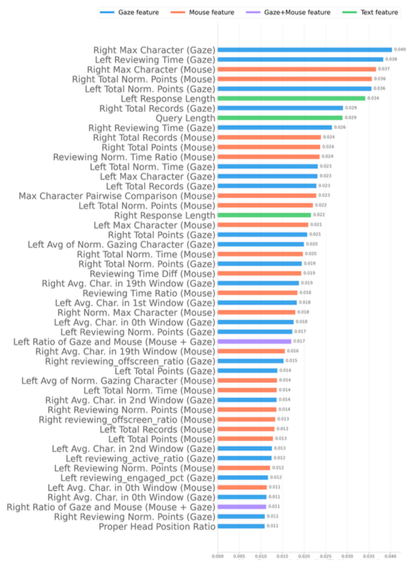



**Figure 5: Distribution of the per-session Pearson correlation between mouse and gaze position, grouped by response length.**




| Reward Source | Short DPO/SFT Score | Short DPO/SFT Win Rate | Medium DPO/SFT Score | Medium DPO/SFT Win Rate | Long DPO/SFT Score | Long DPO/SFT Win Rate |
|---|---|---|---|---|---|---|
| Explicit Feedback | **4.513**/4.385 | 0.494/0.456 | 4.944/4.909 | 0.485/0.476 | 5.456/**5.295** | 0.506/0.455 |
| ModernBERT + Text | 4.447/4.379 | 0.469/0.474 | 4.856/4.830 | 0.481/0.468 | 5.621/5.350 | 0.532/0.425 |
| ModernBERT + Text + IF | 4.594/4.346 | **0.520**/**0.439** | 5.001/4.921 | 0.477/0.476 | 5.599/5.339 | 0.524/0.435 |
| RF + (IF − Gaze) | 4.635/4.417 | 0.481/0.469 | **5.104**/**4.719** | **0.549**/**0.405** | **5.820**/5.388 | **0.576**/**0.385** |
| RF + (IF − Mouse) | 4.492/4.420 | 0.475/0.481 | 4.939/4.947 | 0.466/0.486 | 5.577/5.309 | 0.520/0.432 |
| RF + IF | 4.484/**4.280** | 0.501/0.463 | 5.013/4.880 | 0.486/0.477 | 5.599/5.367 | 0.515/0.438 |

**Table 7: Average quality of the responses with different lengths for 8 LLMs after DPO using different reward models.**

| Method | Accuracy | F1 Class 1 | F1 Class 0 |
|---|---|---|---|
| **Pairwise (695 Samples)** | | | |
| Claude-S-4-6 + Zero-Shot | 0.5303 | 0.5163 | 0.5434 |
| Gemma-4-31B + Zero-Shot | 0.5499 | 0.5252 | 0.5722 |
| mBERT base + Text | $0.5549_{\pm 0.0347}$ | $0.5220_{\pm 0.0366}$ | $0.5815_{\pm 0.0353}$ |
| mBERT base + IF | $0.5526_{\pm 0.0243}$ | $0.3141_{\pm 0.1329}$ | $0.6386_{\pm 0.0247}$ |
| mBERT base + Text + IF | $0.5896_{\pm 0.0193}$ | $0.4733_{\pm 0.1194}$ | $0.6319_{\pm 0.0172}$ |
| mBERT large + Text | $0.5577_{\pm 0.0168}$ | $0.5308_{\pm 0.0194}$ | $0.5810_{\pm 0.0164}$ |
| Qwen3 1.7B + Text | $0.5548_{\pm 0.0261}$ | $0.5101_{\pm 0.0338}$ | $0.5906_{\pm 0.0231}$ |
| RF + All Features | $0.6074_{\pm 0.0073}$ | $0.5776_{\pm 0.0132}$ | $0.6322_{\pm 0.0070}$ |
| RF + (IF − Gaze) | $0.6267_{\pm 0.0097}$ | $0.5935_{\pm 0.0173}$ | $0.6548_{\pm 0.0121}$ |
| RF + (IF − Mouse) | $0.6089_{\pm 0.0130}$ | $0.5836_{\pm 0.0162}$ | $0.6313_{\pm 0.0157}$ |
| RF + IF | $\mathbf{0.6415}_{\pm 0.0114}$ | $\mathbf{0.6084}_{\pm 0.0166}$ | $\mathbf{0.6694}_{\pm 0.0146}$ |
| **Pointwise (248 Samples)** | | | |
| mBERT base + Text | $0.5239_{\pm 0.0667}$ | $0.4716_{\pm 0.1042}$ | $0.5636_{\pm 0.0495}$ |
| mBERT base + Text + IF | $0.5319_{\pm 0.0434}$ | $0.3325_{\pm 0.1903}$ | $0.6254_{\pm 0.0648}$ |
| RF + (IF − Gaze) | $0.6088_{\pm 0.0379}$ | $0.6240_{\pm 0.0432}$ | $0.5846_{\pm 0.0406}$ |
| RF + (IF − Mouse) | $0.6006_{\pm 0.0185}$ | $0.5935_{\pm 0.0252}$ | $0.6044_{\pm 0.0183}$ |
| RF + IF | $\mathbf{0.6291}_{\pm 0.0262}$ | $\mathbf{0.6377}_{\pm 0.0301}$ | $\mathbf{0.6137}_{\pm 0.0297}$ |

**Table 2: Reward model performance comparison (average ± standard error across folds). IF refers to the important features from the implicit feedbacks. RF refers to random forest. Claude-S-4-6 means Claude Sonnet 4.6. mBERT means ModernBERT [48].**

# Conclusions

1. We observe that users tend to look at and point to the responses they prefer. The tendency is so strong💪 that the implicit feedback surpasses classic signals (e.g., the response length, the generation LLMs we use, the LLM judge score, etc) a lot for predicting users' preferences ❗

2. We show that implicit feedback improves reward modeling accuracy from 56% (best w/o implicit feedback) to 64% (best implicit feedback). Moreover, by only using mouse data, the score improvement of SFT + rDPO almost tripled (0.12 → 0.34)!

3. Our extensive analyses show that the mouse trajectories are especially helpful when users need to scroll through long responses. Furthermore, the users, UI layouts, and response lengths all influence interaction patterns.