

Summarization: Contrastive Learning from Exploratory Actions: Leveraging Natural Interactions for Preference Elicitation

Summarized by Xiaonan (Nice) Wang *

Summarization[†]Generated based on Paper from Dennler, Nikolaidis and Matarić (2025)
Topic: Human-Robot Interaction, Multimodal Trajectory Feature Generation, Preference Learning

February 4, 2026

Abstract

This document summarizes the core contributions and methodology of the paper "Contrastive Learning from Exploratory Actions: Leveraging Natural Interactions for Preference Elicitation, Dennler et al. [2025]", focusing on its' main ideas and the core blocks.

1 Overview: Core Questions and Answers

(1) What is the problem?

- Scenario (the Overall Task): **Customize** Robot Behaviors that **Align with Human Preference** -> a.k.a *Preference Learning*
- *Low-Dimensional Preference-Aligned (Semantic) Trajectory Features Generation* using Feature Learning Model (a.k.a Feature Generator)

(2) Why need to solve this problem?

- Generating semantic human preference-aligned trajectory features when customizing a robot, in order to reduce user interactions for eliciting preferences and improve user satisfaction;
- Avoiding data-labelling process like proxy tasks before customizing the robot.

(3) How is it different from prev.?

- Compared with *self-supervised methods* like AEs or VAEs which result in *physically representative features only*, CLEA will learn user preference-aligned features;
- Compared with *hand-crafted* and *prox tasks*:
 - **Complete Feature Spaces** (compared with hand-crafted)
 - **No Extra Labelling Process** (since exploratory search just additionally generates needed data when achieving primary goal)

*wangxiaonannice@gmail.com

[†]**Disclaimer:** This summarization is for research and study purposes. It represents a personal interpretation and may contain inaccuracies. Feedback or corrections via email are highly appreciated.

(4) Why is it better than prev.? (Advantages)

- **Better Semantic Alignment**
- **Lower Human Effort**
- **Trade-off:** Obtained semantic representations without extra human efforts.
- Robot learning as an **exploratory search** (instead of *information retrieval*) interaction is useful for novice robot users (**complete, simple, minimal** and **explainable**)

(5) What is the approach itself?

- **CLEA:**
 - **Contrastive** Learning from **Exploratory** Actions
 - A **Generalizable Trajectory Feature Learning Framework** that provides *Low-Dimensional Feature Vector Generator*

Experimental Workflow as following:

1. Exploratory Search: Exploratory Actions Gen (by User Set A)
2. CLEA Training: Contrastive Feature Generation Learning:
 - Pairwise
 - Triplet Loss
 - Feature Gen Model
3. Downstream Task Evaluation:
 - Behaviors Ranking (By User Set B)
 - Pairwise:
 - For N behaviours in each ranking, building $C(N, 2)$ comparisons;
 - Modeling as $C(N, 2)$ binary classifiers.
 - Why not pointwise?
 - Relativity of Preference
 - Keeping mathematical properties of reward function
 - Non-Binary Classification: Vector Generation rather than Probability Generation
 - Reward Learning (Training) (use 70% pairwise comparisons):
 - Cross-Entropy Loss
 - Reward Model
 - Evaluation (use 30% pairwise comparisons)

(6) What are the applications of it?

- Human-Robot Interaction Field
- A Generalizable Trajectory Feature Learning Framework that can be used in Assistive/Guidance Robots, etc.

2 The Structure

Summarized Block-Diagram The summarized block-diagram of *experimental workflow* see fig. 1¹.

Original in Paper The original block-diagram see fig. 2.

¹For efficiency, this diagram was initially hand-drawn on paper and then converted into its current digital version using Nano Banana Pro.

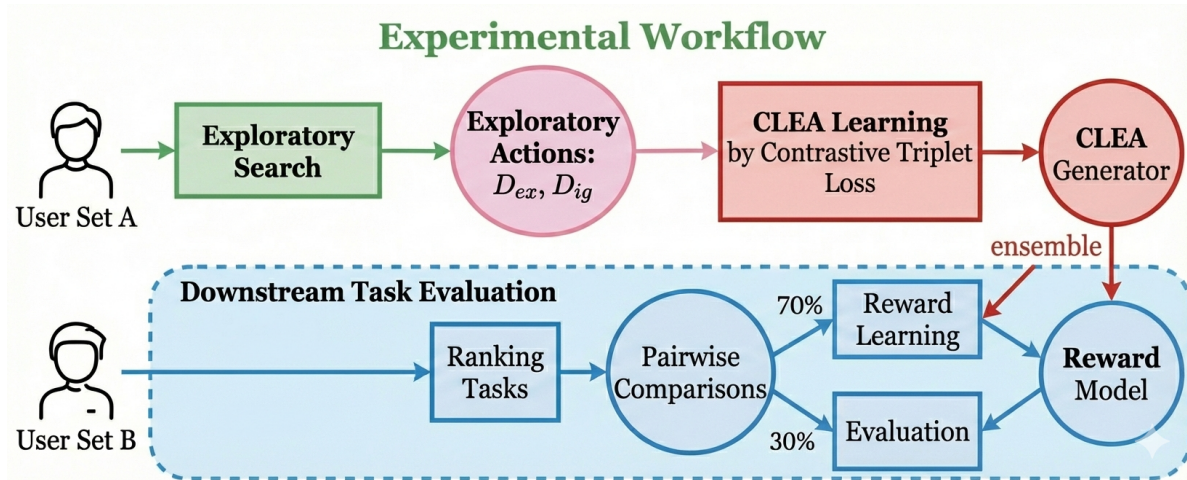


Figure 1: Summarized Block-Diagram of CLEA Experimental Workflow

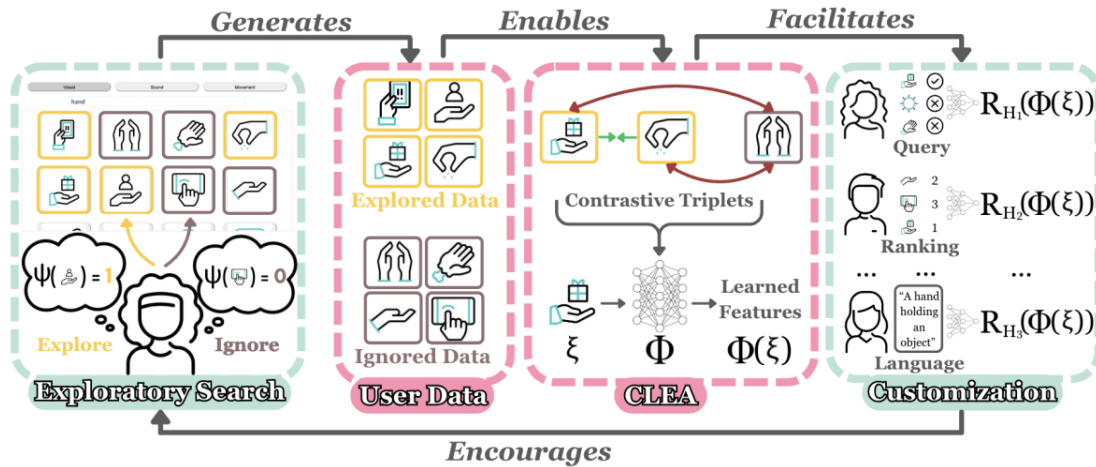


Figure 2: Original Block-Diagram in Paper

3 Open Questions

1. **Purpose of Symmetric Triplet Loss:** Does the second half of Eq. 3 make sense? Will it increase the risk of overfitting while having limited effectiveness in promoting convergence? (In other words, does forcing the “anchor” and “positive” to be interchangeable neglect their differences and lead to “overfitting”?)
2. **Generalizability and Scalability:** Is the trained CLEA feature generator only can be integrated into the robot customization process for online inference to obtain feature vectors? How about other scenarios like utilizing CLEA generator offline to generate vectors to build a feature vector database? If so, does the trained generator meet the generalizability quirements?

References

N. Dennler, S. Nikolaidis, and M. Mataric. Contrastive learning from exploratory actions: Leveraging natural interactions for preference elicitation. In *2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 778–788. IEEE, 2025.