

1. Unsupervised Pre-Training
on General Gameplay Data

Large Dataset
of Human
Gameplay

Pre-Trained
Imitation
Learning Agent

2. Supervised Fine-Tuning
on Task-Specific/
Demonstration Data

Small Task-
Specific or
Demonstration
Dataset

Fine-Tuned
Pre-Trained
Imitation
Learning Agent

3. Reinforcement Learning to
Align Agent with Preferences

Provide
Preferences on
Interactions

Learn Reward
Model from
Preferences

Apply Reward
Model

Interact with
Game
Environment

Aligned
Agent

