Conversational recommendation systems aims to engage with users in real-time, gaining insight into the features of their current needs and then providing recommendations. The challenge in this context is not limited to delivering **accurate recommendation** alone; it also extends to offering **diverse suggestions**. Relying solely on unchanging query and recommendation processes can not only diminish a user's interest in the system but also result in recommendations that consistently favour a small subset of products, thereby also damaging the recommendation accuracy.

Existing conversational recommendation models fall short in terms of diversity, which can be attributed to two primary reasons. Firstly, the majority of them are value-based **single-agent** reinforcement learning models. As long features and generating recommendations are treated as homogeneous actions, controlled by a **single value network**. This value network prioritizes selecting the action with the highest value, which, in turn, hinders the model from exploring other possible actions. Secondly, owing to unbalanced data distribution, the experiential data gathered for reinforcement learning exhibits a **long-tail distribution problem**. A limited number of **popular** items and features dominate the experiential data, causing the model to favor outputs that it has encountered more frequency. This bias towards infrequently seen data compromises diversity.

Considering the two problems, we propose a diversity-enhanced conversational recommendation framework**(DECREC)**. Firstly, independent agents for the two types of actions are established: requesting agent and recommending agent. Additionally, we introduce a **policy-making agent** tasked with analyzing the outputs of the above two agents to reach the final action. This collaborative approach involves **multiple agents** in the action-selecting process, rather that having actions dominated by a single value network. Secondly, **feature entropy** is incorporated into the action value calculation for query features, thereby promoting a more comprehensive range of diverse queries. Diverse queries yield rich feedback, leading to a variety of recommendation pachs and outcomes. Finally, we introduce a **dynamically calibrated experience playback method**. It adapts the frequency of long-tail data and head data to ensure continuous attention to the former.

Extensive experienments prove that DECREC has betteer recommendation diversity than other representative and/or SOTA models.(as the Tabel 1 shows)

TABLE I: Performance of DECREC and baselines. The best results in them are highlighted in bold, and the second-best results are underlined.

Model	Amazon-Book				Yelp				LastFM				Movie			
	SR@15	AT	hDCG	Cov	SR@15	AT	hDCG	Cov	SR@15	AT	hDCG	Cov	SR@15	AT	hDCG	Cov
Abs Greedy	0.204	13.71	0.099	-	0.175	14.58	0.079	-	0.539	11.5	0.231	-	0.719	5.96	0.471	-
Max Entropy	0.447	11.83	0.187	0.295	0.367	12.59	0.129	0.092	0.641	9.90	0.289	0.664	0.694	6.99	0.446	0.146
CRM	0.320	12.30	0.118	0.224	0.168	14.06	0.059	0.047	0.418	12.10	0.141	0.441	0.520	10.57	0.228	0.117
EAR	0.332	12.03	0.127	0.256	0.290	13.11	0.103	0.074	0.578	11.10	0.198	0.576	0.589	8.90	0.293	0.139
SCPR	0.478	11.39	0.199	0.329	0.369	12.50	0.133	0.087	0.681	9.31	0.311	0.624	0.848	4.75	0.536	0.143
UNICORN	0.523	10.90	0.214	0.337	0.380	12.25	0.137	0.090	0.783	8.11	0.314	0.700	0.859	4.31	0.552	0.124
MCMIPL	0.541	10.74	0.239	0.321	0.412	11.73	0.163	0.096	0.832	7.55	0.341	0.708	**0.866**	**3.89**	**0.566**	0.147
DECREC	**0.594**	**9.92**	**0.242**	**0.365**	**0.479**	**11.59**	**0.172**	**0.123**	**0.848**	**7.31**	**0.345**	**0.743**	0.851	4.69	0.539	**0.169**