

Appendix A: Explanation Categories by Item Set Size and LLM

As additional analysis, we studied the explanation categories split by item set size (the number of potential items to recommend to a group). The percentages of attached categories are found in the Table A1. 'Undefined popularity' (Undef. pop.) was mainly found in explanations generated by Llama. However, there was a stark decrease in 'Undefined popularity' when item set size increased (76.2% at 25 items, 23.0% at 75 items). The reverse pattern was found for the category averaging labels (Avg.) and using similarity between users (Sim.). Interestingly, similarity was only found in explanations generated by Llama, not by other LLMs included in this study. All in all, 'Undefined popularity' was not an issue in the explanations generated by Phi, and a minor presence in explanations generated by Mistral and Gemma3.

Patterns were also found for the use of diversity between users (Div.). Both Mistral and Phi increasingly mentioned diversity when item set size increased. In the case of Gemma3, the mentions of diversity remained stable, while we found a decrease in mentions of diversity in explanations generated by Llama at higher item set sizes.

	Avg.			Sim.			Div.			Undef. pop.		
	25	50	75	25	50	75	25	50	75	25	50	75
Llama	21.4	40.6	73.8	26.2	43.2	74.8	14.8	13.0	5.6	76.2	55.8	23.0
Mistral	67.4	86.4	78.4	1.4	4.8	0.8	11.2	35.8	63.8	24.6	10.2	21.6
Gemma	74.2	83.6	79.2	4.8	0.6	1.2	26.8	23.2	26.0	23.8	15.8	18.0
Phi	97.2	98.6	94.4	0.2	0.2	0.8	28.0	30.4	46.2	2.2	1.4	6.4

Table A1: Percentages of categories mentioned in explanations; by LLM (Llama, Mistral, Gemma and Phi) and by Item set size (ranging from 25 to 75 items). Included categories are: (i) *averaging ratings* (Avg.), (ii) *similarity between items/users* (Sim.), (iii) *diversity in recommendation list* (Div.), and (iv) *using undefined popularity metric/threshold* (Undef. Pop.)

Additional exploratory analyses between uniform and divergent groups did not lead to further insights. No differences were found between the two group configurations.