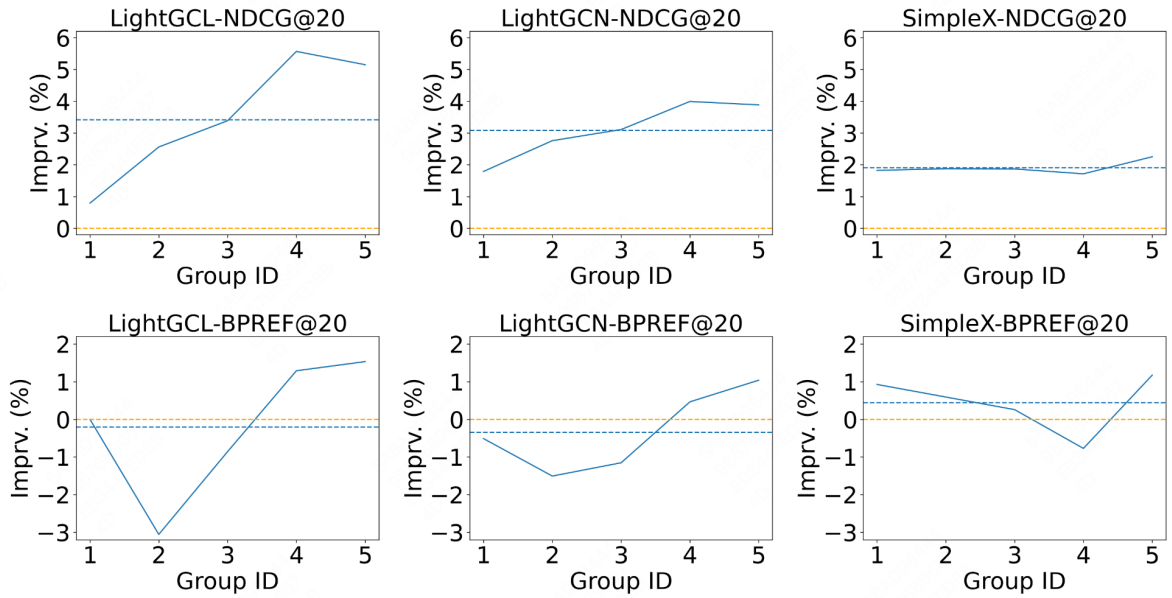

Figure 4: The improvement rates of 5 user groups in 3 backbone models on Amazon-CDs and Vinyl datasets. Blue dashed line denotes the average improvement rate of 5 groups, and orange dashed line denotes zero value.

In order to study the impact of the richness of user history on the recommendation results, we deliberately sampled the dataset, making it consist of 5 user groups with different richness of histories. We take 3 models that perform relatively well in the Amazon-CDs and Vinyl dataset, and calculate the improvement rates of NDCG and BPREF on 5 user groups, as are shown in Figure 4. It can be discovered that generally users having richer rating histories can achieve better improvement rates after applying UPLLM. This shows that UPLLM is especially capable for long history modeling, which is of great benefit when applied in real world scenarios.