

預測式資料增強稀疏圖神經網路推薦系統

Predictive Data Augmentation for Sparse Graph Neural Networks Recommenders

指導教授：李政德

專題成員：李秉儒

開發工具：python

測試環境：window 10

一、簡介：

基於增強稀疏圖神經網路的想法，我們對不同的推薦模型進行多種資料集的測試，以觀察不同模型和資料集對於該方法的適應性和增強性。

測試方法主要有兩種：

1. 將模型預測的結果(top20)重複加入到原本的訓練集中，其中只加入預測錯誤的部分(移除和測試集中重複的部分)，以達到提高預測率的結果。
2. 減少原先訓練集的資料以獲得更為稀疏的資料，使用該資料訓練模型後並從中分析不同稀疏資料集的表現。

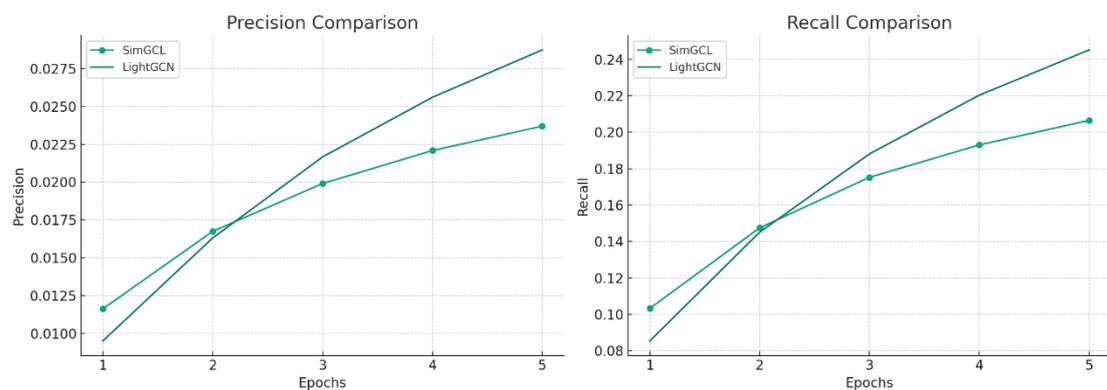
測試的模型：

1. Diffusion Recommender Model
2. LightGCN
3. SimGCL(XSimGCL)
4. LLMRec
5. UltraGCN

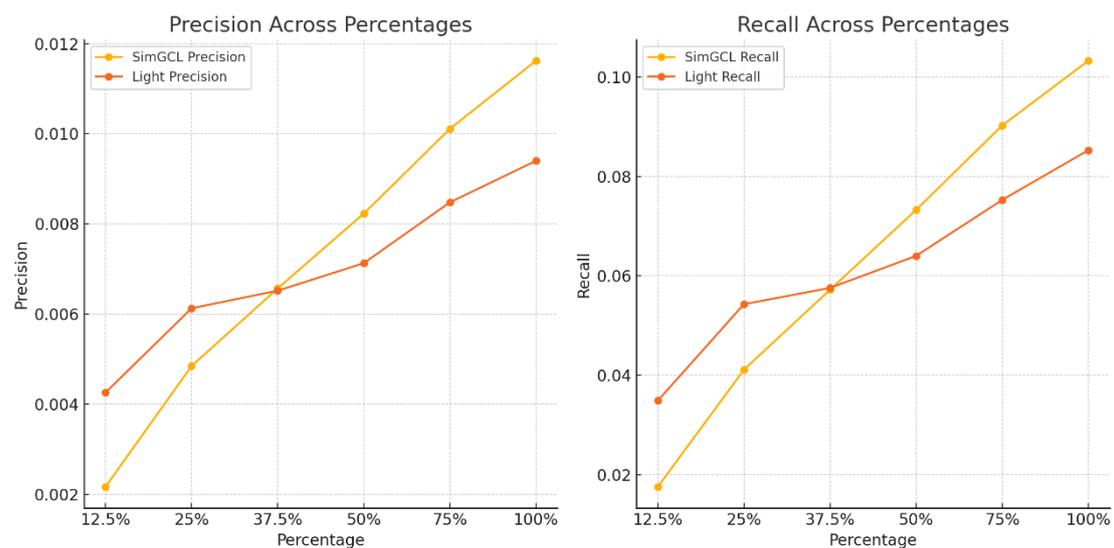
測試資料集：

1. Yelp2018
2. Netflix
3. Gowalla

二、測試結果：



上圖為 SimGCL 和 LightGCN 在 Yelp 的資料集中的表現，其中 Epochs 代表加入過幾次 Top20k 的結果(1為原始資料)。



上圖則為縮減資料集後訓練的結果(Top20)，模型和資料集一樣為 LightGCN，SimGCL 和 Yelp。