Team: Turing Twins

Problem Statement 2

Problem trying to ideate:

Recommender System: Preference predicting system, based on historical data. Currently using it for predicting co-purchased products for market feasibility.

Recommend Books from co-purchased books.

Recommend Videos from co-purchased Videos.

Recommend Music-CD from co-purchased Music-CD.
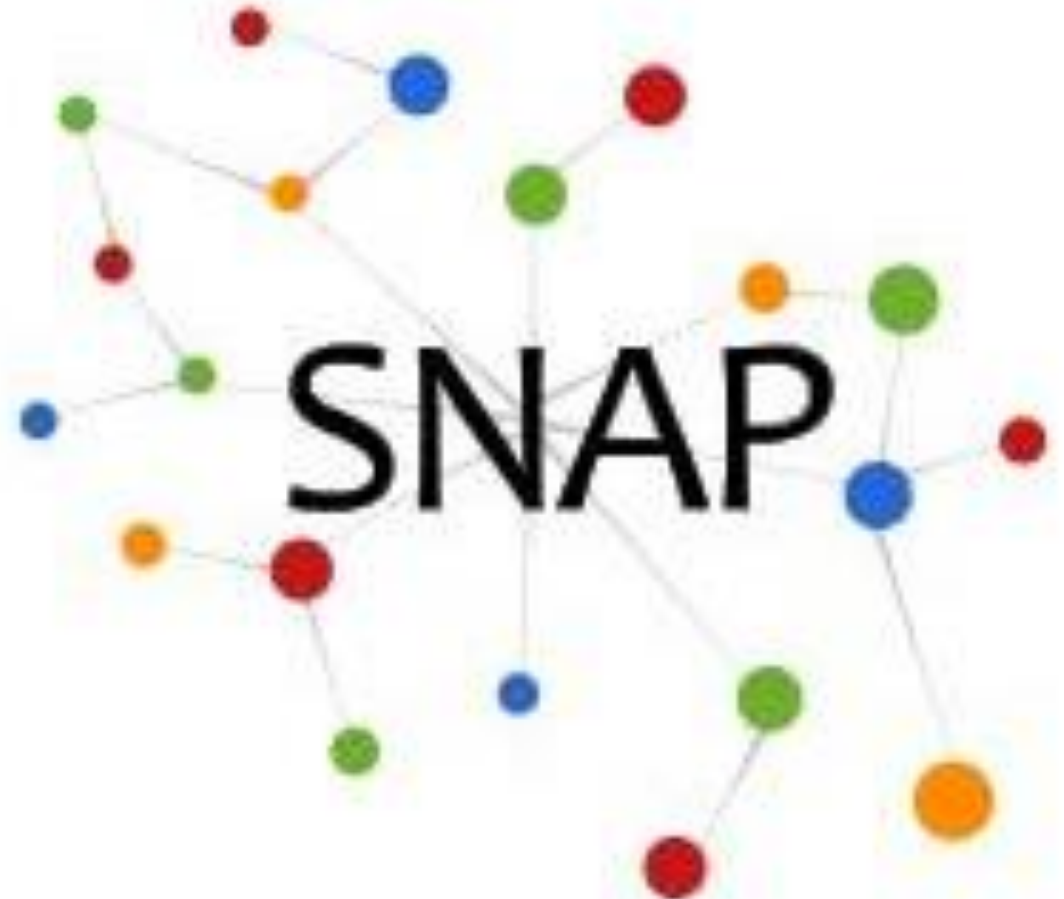
Recommend DVD from co-purchased DVD.

Created a SNAP Recommender system for all items in the whole dataset.

Use Case Identified:

- Pre-processing Amazon SNAP user data and meta data.

- Predict co-purchases by the customer as to enhance buying and selling of goods on retail websites like Amazon. Also recommend items for the purchases made.

Tech Stack

- Nltk for pre-processing

- Regex for feature extraction

- NetworkX for plotting, visualizing and feature extraction.

- Packages like numpy, pandas etc

- Girvan Newman Algorithm for community prediction on frequently purchased items.

Solution Approach:

- We are trying to solve the problem associated with understanding the huge Amazon SNAP dataset and thereby pre-processing it to recommend the co-purchased items.

- We have a two way solution:

- Recommendations based on community predictions using directed subgraphs of SNAP dataset with Transaction data and Meta Data, using Girvan Newman Algorithm. Now we have directed graphs, using which top and next level communities are predicted.

- So when Amazon uploads transaction data of any time frame, using our approach they can predict the frequently co-purchased items and focus more on highlighting those items to the target customer.

- For a more detailed and in depth analysis and recommendations, recommendation systems for individual items. The meta data file contains 4 features for which we are able to predict the co-purchased items using graph based approach and analytics.

Roadmap

- Data-preprocessing. Using NLP and regex.

- Creating 4 different graphs, txt and edge lists for Books, DVD, Videos, Music-CD by pre-processing them individually.

- Created individual recommendation system for the four categories by extracting features from graph.

- Joining the user data with the meta data.

- Created a recommendation system using Girvan Newman Algorithm for suggesting co-purchased items on the whole data set. Applied using community detection.

Architecture Diagram
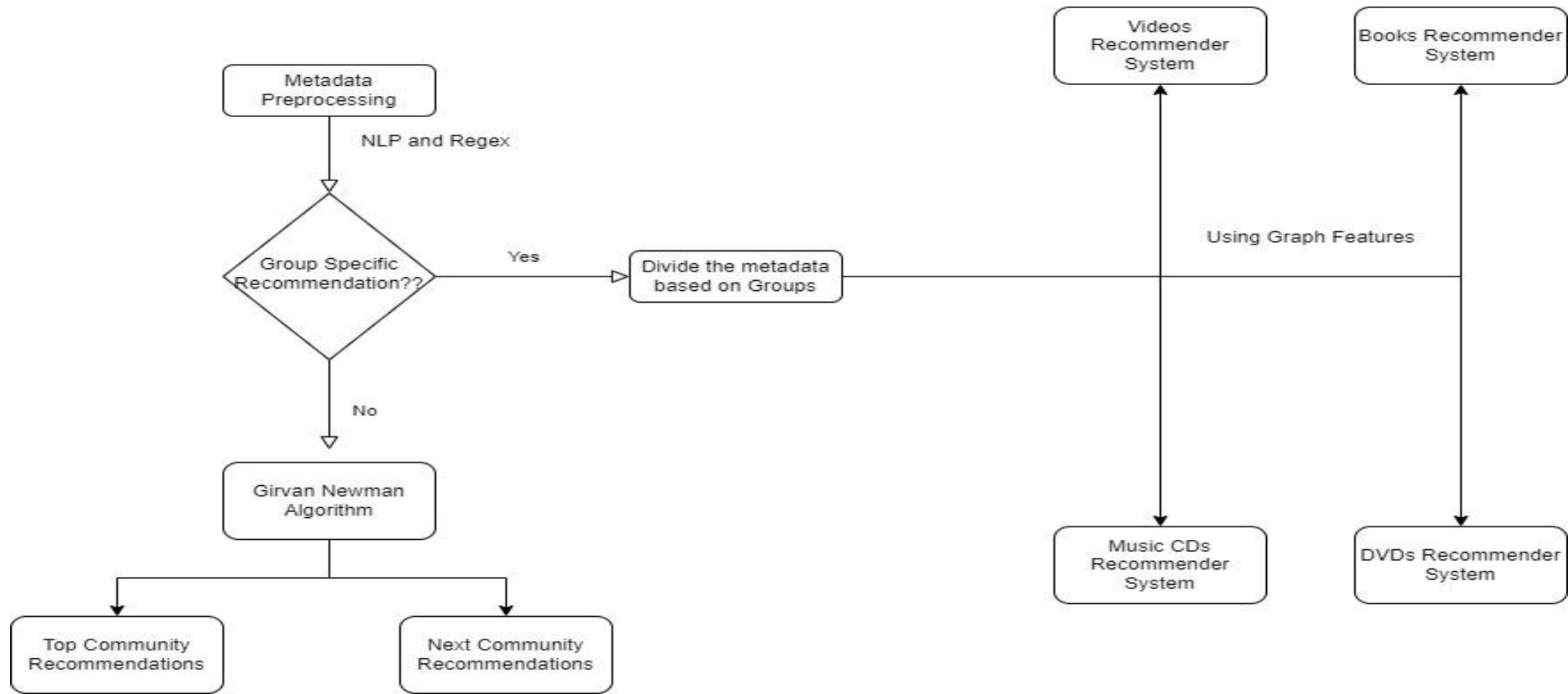
Why use us?

- We solve the problem of predicting similar items using highly efficient Girvan Newman Algorithm. We predict the co-purchased items based on top level (most co-purchased) and next level(on a second level that is distant co-purchases).

- We give a deeper solution as to if we have a customer wanting only to purchase books or only DVD or only videos or Music-CD. This feature can be enabled using our separate recommendation system as well.

- Thus our USP: **One for all and all for one.**

Future Scope

Build and deploy a user friendly application for our recommendation systems.
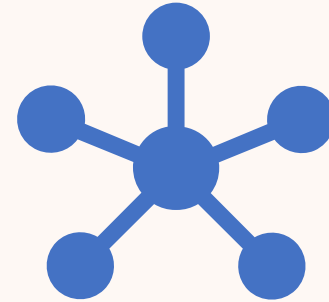
Use Graph Neural Networks for better results.