



LASTFM Users Country Prediction

Team: Alexander Shirnin, Nikita Markin, Alexander Sukharkov



Requirements:

- GitHub project: https://github.com/25icecreamflavors/SNA_2023
- Communication channel:
https://github.com/25icecreamflavors/SNA_2023/discussions
- Project team: Alexander Shirnin, Nikita Markin, Alexander Sukharkov



Project idea

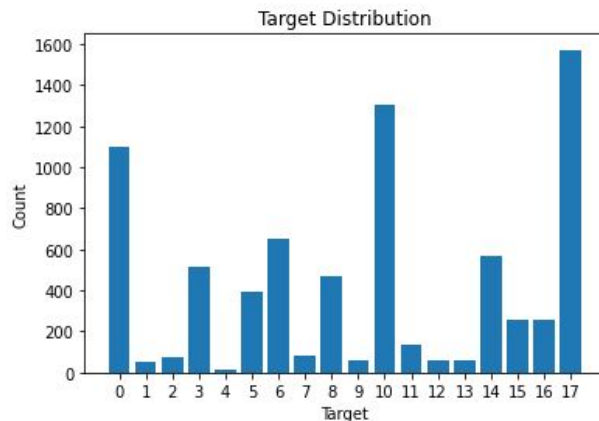
- In this research project, we aim to predict the country of Asian Last.fm users based on the users they follow and the musicians they like. We will use machine learning and graph-based deep learning methods.
- We are going to use different methods for nodes representation and compare model performance on this dataset

Steps to achieve the goal

1. Use Node2Vec and DeepWalk to create nodes embeddings.
2. Use XGBoost and LogReg trained on embeddings to predict the target country.
3. Use GAT trained on embeddings (from part 1 or from linear layer) to predict the target country.
4. Compare the results.

The dataset

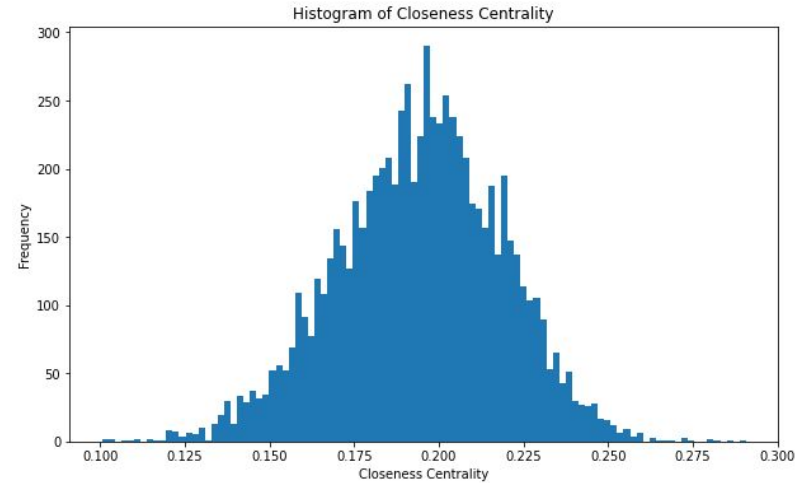
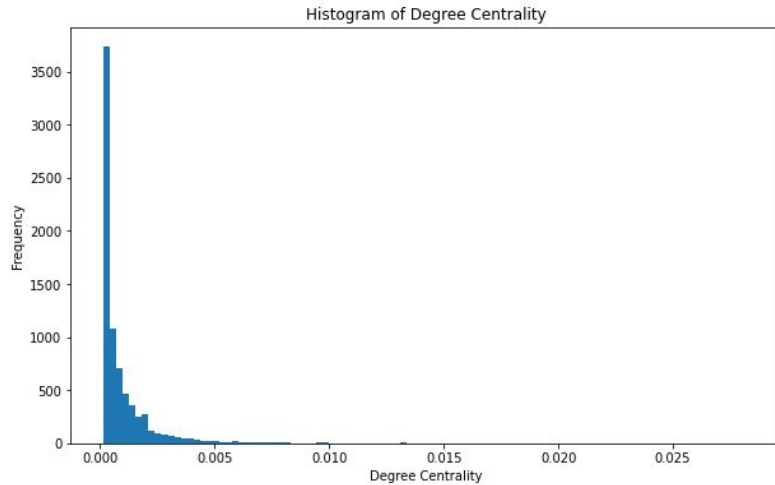
- Dataset used from: [LastFM Asia Social Network](#) (Snap Stanford Datasets)
- 7,624 nodes (users in total)
- 27,806 edges (users following each other)
- Each node has a label (country). The target feature was derived from the country field for each user profile, might contain noise.
- In total there are 18 countries.



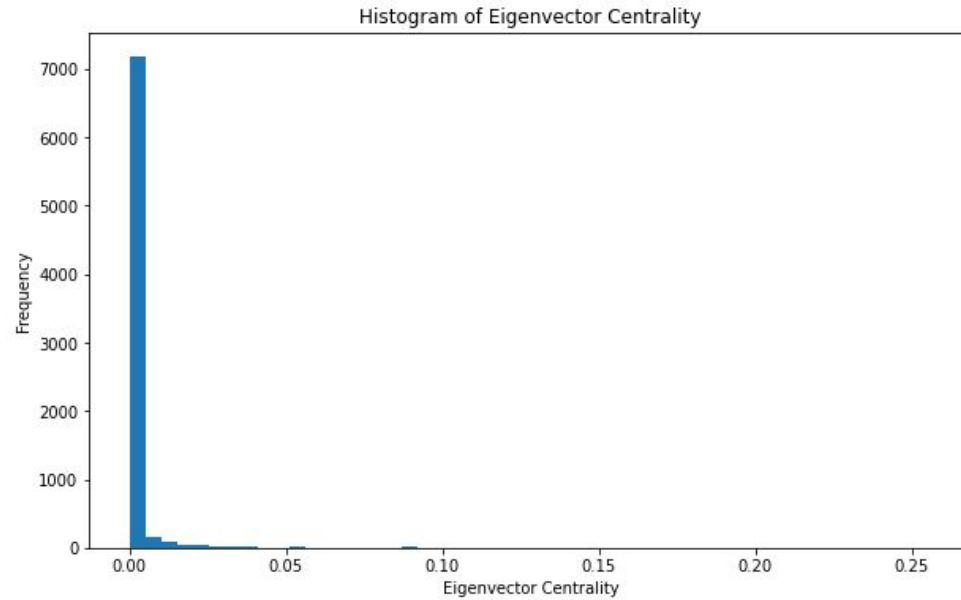
Network description

Diameter: 15

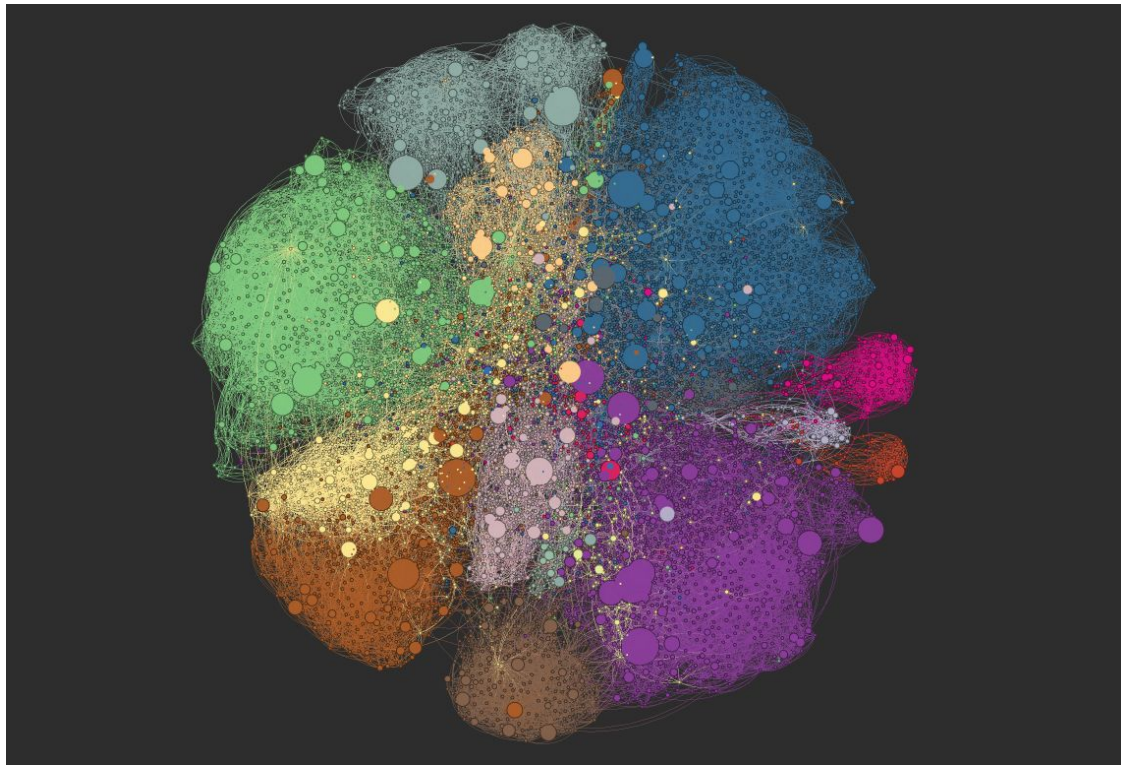
Density: 0.00095



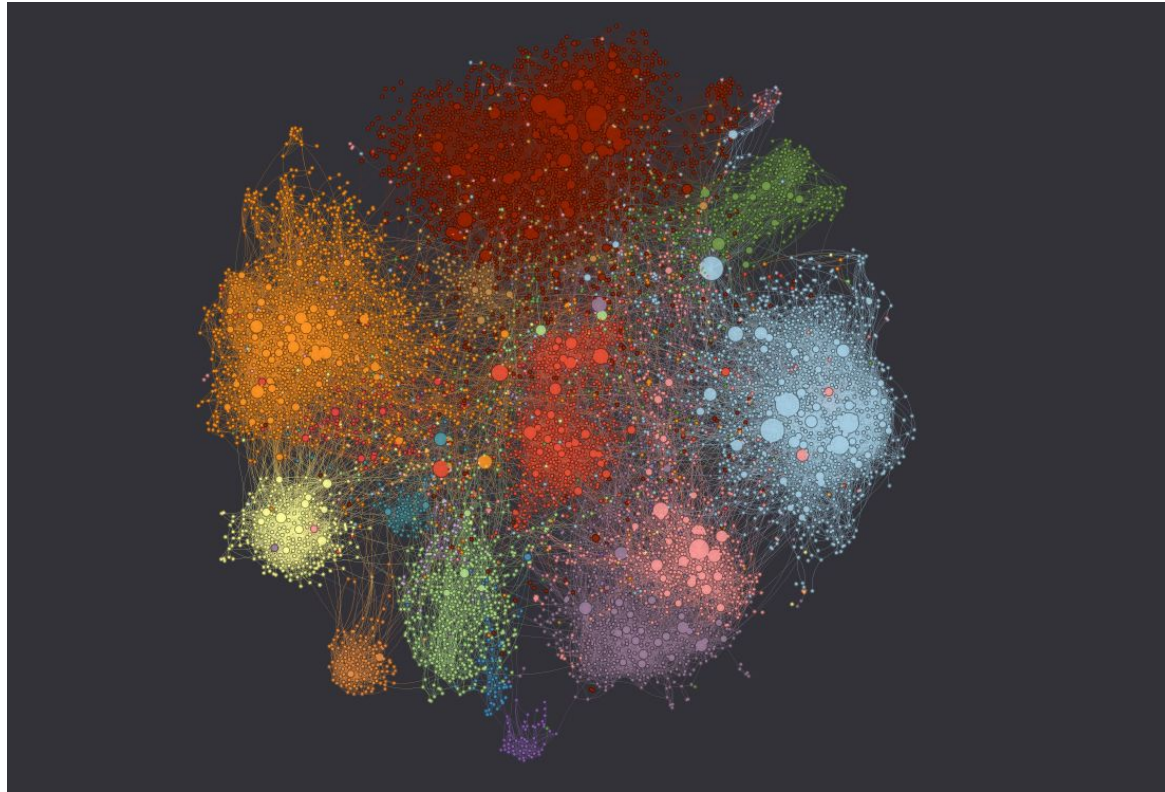
Other metrics



PageRank



Betweenness centrality



That's all.