

This folder is contained selected files from my M.Sc. Thesis. In this study, the interplay between Instagram communication network and Homophily with a focus on visual contents is studied. To achieve this, a mobile application was developed that helps Instagram users to download their favorite images through it. Users were divided into three groups namely Close-Friend(based on their like and comments and link), Link-Friend(base on social links) and Not-Friend(randomly). For each group, high-level (DeepSentiBank was compiled and used for sentiment analysis) and low-level (matlab libraries) features from the images they were interested in were extracted. To measure the similarity, two novel methods for profiling users based on high-level features and one distance measure method based on low-level features were proposed.

First, we confronted the huge, dirty data that we cleaned it by variery methods. Then, for all remaining users we used DeepSentiBank for analysis images and assign score to each predefined tags. After that we implemented **lable.py** to make high level vector for each user and **high.m** for calculating similarity between two users by means of their high level vectors. For low-level features as I mentioned, I used Matlab predefined functions for extracting each image's low level features then I implemented **low.py** for measuring the similarity between two users by analyzing analogy between all images of two heads of an edge. For testing I used **CreateRandomLinkFriend.py** and **CreateRandomNotFriend.py** for creating random users and also used plot.py to show the results. I also suggested a modified algorithm in **modify.py** to improve the DeepSentiBank assigned tags. As I mentioned this folder is just a part of my project.