Aili Harmon
Daniel Shu
Zack Shatsky

Final Project Proposal

We are interested in building a predictive model for users of social apps focused on weight loss. Our primary goal is to construct a Graph Convolutional Network for regression that uses user characteristics, such as age, gender, and BMI in addition to user interactions such as follows and comments to predict the total weight loss of users over the duration of their activity on the app. We would also like to examine the relative effects of different user characteristics and interactions on the strength of our model and plan on training multiple versions of the model that exclude the use of different features. Our evaluations will be based on the mean squared error of our models on a held-out test set. Finally, we may also investigate other deep learning models. Graph Recurrent Neural Networks could be used to take advantage of the different networks that could be formed based on single-user interaction types. A standard fully connected neural network that takes inputs with network information summarized and encoded as features (e.g. for each user, use their degrees as features) could be interesting to explore. For these comparisons, we would evaluate both training time and test error.

We plan on using the weight loss application data set provided by Professor Derr, which was collected from a popular social app for weight loss in China. Because of the massive size, we will first filter out less useful data (e.g. users with very low degrees). Since we will also be training models with different features, we will need to manipulate the data into the correct form for each of them. For training and testing, we plan on using an 80-20 split.

Timeline

Week of 10/24: Write proposal and obtain the data set
Week of 10/31: Perform data preprocessing and begin building preliminary GCN model
Week of 11/7: Gather preliminary results (MSE) for project update
Week of 11/14: Collect results on various feature exclusions
Week of 11/21: Thanksgiving
Week of 11/28: Improve existing GCN model or explore other models
Week of 12/5: Finish experiments and prepare for presentation
Week of 12/12: Finish writing paper