

# Collaborative and Content Based Approaches to Recommendation Systems

By: Victoria Schaller and Claudia Levi

## Problem Statement and Data

### Problem Statement

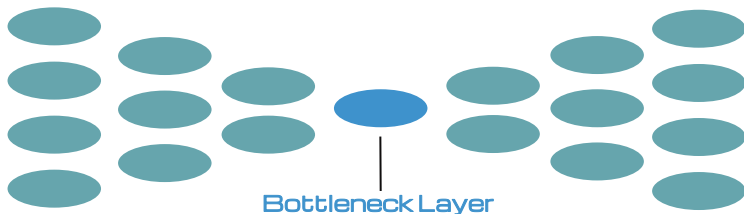
- Over 100,000 board games, searching for a game can feel exhausting
- Existing platforms rely on general popularity of manual filters

### Data

- 22,000 games and 411,000 users with 19 million ratings
- Over 4000 features including name, description, # of players, rating distribution, themes, mechanics, subcategories, artist information, designer information, and publisher information

## Content Based Filtering with Denoising Autoencoder

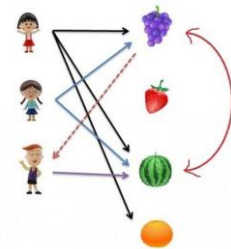
- Denoising autoencoder to embed game features into a 64 dimension space
- Content based filtering to leverage the embeddings to predict how users would rate games



## Item-Item Collaborative Filtering

- User Inputs: # of players, game category, game sub-category, games they enjoy, games they don't enjoy
- Create User-Item matrix
- Create Train/Test Split for each user (80/20)
- Calculate cosine similarities between games
- Predict ratings of candidate games using a weighted average
- Return Top 5 Games

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|}$$



## Results

- Collaborative filtering had an RMSE of 1.3 and a slight tendency to overpredict compared to underpredict Hybrid w/Autoencoder
- Denoising autoencoder captured useful latent game features that the content filtering leveraged to predict user ratings. Despite some regression to the mean, rank order of game ratings was preserved

