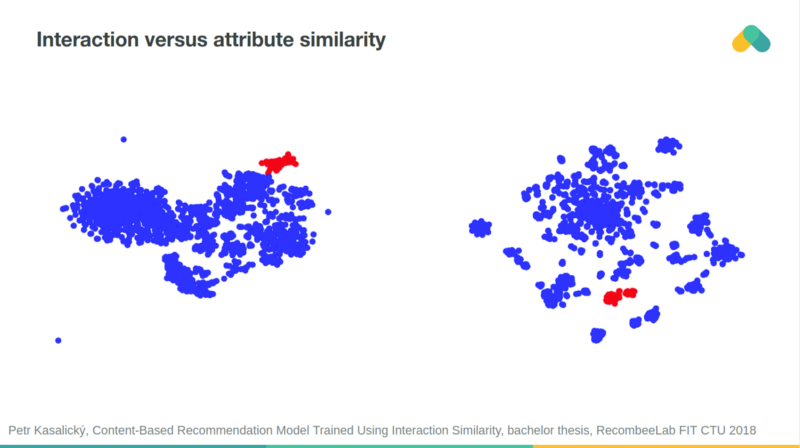
Data collection

This is the first and most crucial step for building a recommendation engine. The data can be collected by two means: explicitly and implicitly. Explicit data is information that is provided intentionally, i.e. input from the users such as movie ratings. Implicit data is information that is not provided intentionally but gathered from available data streams like search history, clicks, order history, etc.

Initially, with the help of google form, we can collect the primary data from different people of different age groups, for the data set explicitly. Then the data set can be splitted into training and testing set. Splitting the data into training and testing sets is an important part of evaluating predictive and recommendation modeling. The idea of it is to train the recommendation system with the larger portion of the data for training , which may include certain fields (such as age, gender, destinations, etc.) and, accordingly with the help of testing set of data we could be able to test the machine for its accuracy and performance. Since this project has a cold start problem it can be reduced when attribute similarity is taken into account. Cold start problem is not having enough historical interactions. You can encode attributes into binary vector and feed it to recommender. 

Items clustered based on their interaction similarity and attribute similarity are often aligned. You can use neural network to predict interaction similarity from attributes similarity and vice versa.

