Description of the required setup

**Required packages:**

Anaconda above 3.0

**Steps:**

1. Download the RecSys 2015 challenge dataset from: https://www.kaggle.com/chadgostopp/recsys-challenge-2015
2. Run demo.py, input the path that dataset are located. Typing one of the 2 commands (preprocess/analysis)
3. Type preprocess commands to get following documents into dataset folder:
4. yoochoose-clickssave.txt: each row represents a user, columns are: user id, start week, start hour, end week, end hour, dwell time, number of clicks, number of items, label
5. yoochoose-clickssave1.txt: randomly shuffle yoochoose-clickssave.txt for Predict Potential Customers by using neural network and KNN classifier
6. yoochoose-transaction-clicks.txt: for all users, each row represents a user, the first element is user id, the others are this user’s unique selected products (USP)
7. yoochoose-transaction.txt: the subset of yoochoose-transaction-clicks.txt which includes users only with bought behavior
8. yoochoose-sequence-clicks.txt: for all users, each row represents a user, the first element is user id, the others are this user’s actual selected products (ASP)
9. Type analysis command to predict potential users and their interesting purchasing items. Run to get finalsolution.txt in the dataset folder. Each line in this file include the user id followed by the predicted items.