"Bandit with Knapsack" Algorithm

Michael White and Matthew Wolfson
INFO 4099, Cornell University
Correspondences: mhw66@cornell.edu, mcw88@cornell.edu

I. Initial Goals of the Algorithm

Generate a list of activities from a given set of a data (past exercises - type, length, calories burned, time of day completed) that slowly increases a user's activity towards a goal of 10% overall growth.

II. Approach

The approach towards completing this task was to use the ideas of a known algorithm called "Knapsack" to find the best combination of exercises based off of payoffs for each activity, i.e. energy expenditure, and increasing the size of the knapsack a little each day. To make the creation of the list more intelligent we made adjustments to payoffs to take into account the user's adherence to the activity, days past since last completed, and how often it has been completed (since activities completed more often are considered "easy-to-do" as the user is already good at it). In addition, we used the ideas of exploration vs exploitation that is found in the "Bandit" problem. In terms of exploration the algorithm, with a probability of 5%, randomly removes the bottom two activities from the knapsack list and replaces them with two random activities from the set. The exploit part on the other hand is just that, the calculated knapsack list is the outputted list – i.e. no exploration occurs with an exception; the knapsack is updated every day taking into account all of the previous day's activities (new data) and ensures that the outputted list is still optimal.

III. Implementation

The algorithm was implemented in python (see attached file) and tested with the data we were given. Note, however, that we made the algorithm to take input in time order, meaning the newest added activity was the most recently completed. With this, upon a user completing an activity it can be added to the data set and it does not affect the current outputted list as a new list is only generated once a day.

IV. FUTURE IMPROVEMENTS

The Biggest change that we need to take into account is what we are trying to increase. From what we learned from the physios, we need to work toward getting a user's daily activity to a total of 150

minutes a day, instead of an overall 10% increase. In addition, activities that are easier to do (walking compared to running) should be suggested more often as they are better for the user.

Another improvement that can be made is to figure out a way to pick items from the knapsack to be displayed to the user based on most probable time of completing. We discussed this, however we are not sure what the final decision was regarding it.