**Problem One**

A) Feature set <number of friends, number of photos, number of videos, number of statuses>

This feature set consists of a variety of indicative features that can be measured using real numbers. Number of friends, number of photos, number of posts, and number of videos can all be measured using any integer greater than or equal to 0. This feature set is useful is gauging what type of Facebook user it is as it clearly represents user activity and engagement.

B) Given the above feature set, a metric could be the total number of likes the profile has. This measurable could be clearly compared across profiles, and would be a good indicator of how popular a given user is. If a user receives more likes than another user when taking into account pictures, videos, and posts, then this is a holistic view of general online popularity.

C) No it does not make sense to put people in a Euclidean 3-space. I would place the most weight on number of hairs on head. This metric is not dependent on gender, because no matter what it is likely that a child has more hair than an elderly person; this trend is clear across gender lines. That being said, the distribution would definitely be more extreme for the male gender. I would then weigh height and weight at the next tier, as they are less indicative of age. While it is true a shorter, lighter person is probably a child, this is not as clear of a pattern and is a bit more difficult to discern across gender.

D) As explained in the paper, one could assess the distance between strings based on how many edits are necessary to make them identical. The three different means to edit a given string are insertion, deletion, and replacement. Weighted equally, these methods can determine the distance between two strings by counting how many times each needs to be applied. This metric could also be used to discern the distance between two DNA strands. Because DNA strands are encoded with 4 bases, each represented with a letter, their representations take the same format as a standard string. Because of this, the method described in the paper could be used as it stands to determine the distance between two DNA strands. In order to see how two DNA strands’ bases compare, this algorithm could assess the distance between their two string representations.