**Heuristic Algorithm for exercise 1:**

1. **Given:**
   1. The provided data-set for this task contains data for two classes of data: ‘Trump’ and ‘Hillary’.
   2. The most informative features are (from analysis from exercise 1) :
      * retweet\_count
      * favorite\_count
      * text\_len – most of the Trump tweets are longer.
      * source\_url – most of the Trump tweets come from android, and Hillary from tweeterdeck.
      * handle – class label.
   3. Unlimited space , but the algorithm should be optimal in computation complexity, since computation resources limited .
2. **Methods**:
   1. **Method 1**
      * Creation of rules , such as predefined classification tree, which based on features (analyzed in solution 1). This solution would work since all features are informative, for example: most of the tweets that ‘Trump’ produce from Android device, have longer texts, more favorite\_counts and retweet\_counts.
   2. **Method 2:** 
      * Define a cost function for the cost of an edge between each 2 nodes.
      * Create a graph for each class, and .
      * On each graph execute calculation of [Modularity](https://en.wikipedia.org/wiki/Louvain_modularity).
      * According to computation limitation and accuracy find number of hub nodes (based on communities) and store them as leading nodes of sub graph (neighbors in communities).
      * Prepare a list of
      * ***Classification***:
        1. get new record.
        2. In each graph get the most closest by cost function or **Method 1**.
        3. In each of the 2 selected communities (by the hubs) perform local search on random and limited number of nodes, and select the best one (in each graph).
        4. Between 2 nodes, from and graph select the best one as in step 2.
        5. return winning class.
      * This solution very fast in run time for classification and pretty accurate, doesn’t need a lot of computation power.