

## Theoretical task 3.

*Recommendations: all solutions should be short, mathematically strict (unless qualitative explanation is needed), precise with respect to the stated question and clearly written.*

1. Prove that the complexity (number of elementary mathematical operations such as  $+$ ,  $-$ ,  $*$ ,  $/$  for scalars and boolean condition checks) for binary decision tree training from training set with  $N$  objects having  $D$  features each doesn't exceed  $O(DN^2 \log_2 N)$ . At this task suppose that the decision tree is balanced and each terminal node contains only one object.
2. Assume we use a particular feature  $i$  for splitting a certain internal node containing  $N$  objects. We want to determine the optimal threshold value  $h$  such that the sum of entropies for the children nodes  $\{x|x_i < h\}$  and  $\{x|x_i > h\}$  is the smallest. Propose the algorithm that finds the optimal threshold with complexity  $O(N \log N)$ .
3. Consider multiclass classification task with  $K$  classes. Objects in this task have  $n$  binary features. How many different decision trees of depth  $n$  can be constructed for this task? Two decision trees are different if there is an object in  $\{0, 1\}^n$  which they classify differently.