

# MIW (eng) - mini-project 2 (1.04.20)

## Classification - kNN

1. Download any dataset you like from UC Irvine Machine Learning Repository for the classification purposes:

<https://archive.ics.uci.edu/ml/datasets.php?format=&task=cla&att=&area=&numAtt=&numIns=&type=&sort=nameUp&view=table>

Identify which columns you'll be using as features and which for labels.

2. Implement kNN with 2 or more dimensions on your chosen dataset:
  - a. Split the dataset into train/test datasets.
  - b. Run the algorithm.
  - c. Don't use Euclidean distance, use other metrics (e.g. Manhattan).
  - d. Evaluate the algorithm's correctness (predicted correctly over total).
  - e. Plot the results.

You can use libraries and modules like numpy, sklearn, sklearn.preprocessing, plotka, matplotlib, pandas as helpers. For example, dot product, label encoding, opening a file, showing statistics about the file (pandas). I want to see that you are writing code by analyzing what you are doing.

## Deadline rules:

1. By 01.04.20 - max amount of points is 10
2. By 08.04.20 - max amount of points is 10
3. By 15.04.20 - max amount of points is 5
4. Later - 0 points