

HOMEWORK 1 DOCUMENTATION

Authors: Manuel Di Pietro and Lorenzo Pichilli

This homework is implemented using the **Python 3** programming language. Our algorithm is based on the **K-nearest neighbors algorithm** and it works like that: for each query in the query set we obtain the current query and we loop into our training set to find a list of the nearest neighbors limited to **K** elements. To find every possible neighbor we calculate the distance between the current training element and the current query, for example:

```
Current training element: 357, 389, 459, 685, 748
```

```
Current query element:   350, 379, 359, 665, ?
```

The distance is calculated like that:

$$|357 - 350| + |389 - 379| + |459 - 359| + |685 - 665| = 137$$

After performing all the calculations and finding the **K**-nearest elements, to predict the possible value that will be assigned to the “?” marker in the current query, we calculate it using the weighted average among all of the **K**-nearest elements found. The more the distance is near to 0 (distance = 0 means that the query is equal to the training element) the more its value is taken into account in the weighted average calculation.

In the end, every query, with marker replaced, will be written into the result file.