

G V V Sharma\*

## CONTENTS

1	Problem Statement	1
2	Algorithm	1
3	Examples	1

**Abstract**—This manual introduces the nearest neighbour method.

## 1 PROBLEM STATEMENT

- 1.1 Let  $\{x_i\}_{i=0}^{N-1}$  be the marks obtained by students in a class. The objective is to award grades ranging from  $A$  (highest) to  $D$  (lowest) to the students. This is done by mapping a set of marks (which are closer to each other in some sense) to a cluster.
- 1.2 Table 1.2 shows the equivalence of each grade to a cluster. Thus, the objective is to map  $x_i$  to

Cluster	$C_6$	$C_5$	$C_4$	$C_3$	$C_2$	$C_1$	$C_0$
Grade	A	A-	B	B-	C	C-	D

TABLE 1.2

$C_j$ .

## 2 ALGORITHM

- 2.1 Let  $m$  be the iteration number.
- 2.2 **Initialization:**  $m = 0$ . Compute the mean value of the cluster  $C_j$  as

$$\mu_j^0 = \frac{j}{K-1} \max_{0 \leq i \leq N-1} x_i \quad (2.1)$$

2.3 **Fitting:** For  $x_i$ , let

$$|\mu_k^m - x_i| < |\mu_j^m - x_i| \forall j \quad (2.2)$$

Assign  $x_i$  to the cluster  $C_k$ .

- 2.4 Let  $x_{i,k}$  be the set of all  $x_i$  assigned to cluster  $C_k$  in the  $m$ th iteration. Let  $N_k$  be the number of such elements. Revise the mean value of  $C_k$  as

$$\mu_k^m = \frac{1}{N_k} \sum x_{i,k} \quad (2.3)$$

- 2.5 Repeat the above process till  $\mu_k^m = \mu_k^{m+1}$ , i.e. the algorithm converges.

## 3 EXAMPLES

- 3.1 Demonstrate the nearest neighbour algorithm through a python script.

**Solution:** Execute the following

```
svn checkout https://github.com/gadepall/
EE1390/trunk/manuals/nearest_neighbour
/codes
python grades.py marks.xlsx
```

The **grade\_new.xlsx** file will contain the grades.

- 3.2 Modify the python script to add an  $F$  grade as well.

\*The author is with the Department of Electrical Engineering, Indian Institute of Technology, Hyderabad 502285 India e-mail: gadepall@iith.ac.in. All content in this manual is released under GNU GPL. Free and open source.