

Tables and CD diagram with a statistical comparison of the AUC results of our proposal using  $k=25\%$  and other contrast pattern selection methods reported in the state-of-the-art, considering all the tested databases.

Octavio Loyola-González  
octavioloyola@bioplantitas.cu

José Fco. Martínez-Trinidad      Jesús Ariel Carrasco-Ochoa  
fmartine@inaoep.mx              ariel@inaoep.mx

Milton García-Borroto mgarciab@ceis.cujae.edu.cu

February 17, 2017

#### Abstract

In this document we show supplementary material for the paper entitled “A Novel Contrast Pattern Selection Method for Class Imbalance Problems” submitted to the 9<sup>th</sup> Mexican Conference on Pattern Recognition (MCPR2017).

## 1 Average rankings of Friedman test

Average ranks obtained by each method in the Friedman test.

Algorithm	Ranking
BestCP	6.9789
Covering	4.8
Best K 10	4.4
Best K 50	2.8158
Best K 80	2.7684
Proposal 10	3.0684
All CP	3.1684

Table 1: Average Rankings of the algorithms (Friedman)

Friedman statistic considering reduction performance (distributed according to chi-square with 6 degrees of freedom: 288.107143.

P-value computed by Friedman Test: 1.1006562328219616E-10.

## 2 Post hoc comparison (Friedman)

### 2.1 CD Diagram

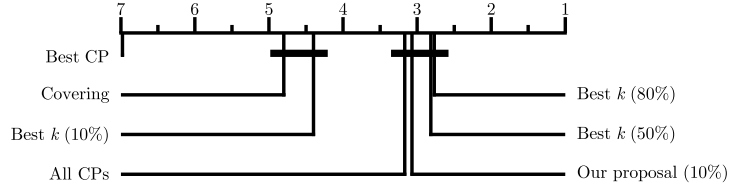


Figure 1: CD diagram with a statistical comparison (using  $\alpha = 0.05$ ) of the AUC results of our proposal using  $k=25\%$  and other contrast pattern selection methods reported in the state-of-the-art, over all the tested databases.

P-values obtained in by applying post hoc methods over the results of Friedman procedure.

$i$	algorithm	$z = (R_0 - R_i)/SE$	$p$	Finner
6	BestCP	13.43321	0	0.008512
5	Covering	6.481524	0	0.016952
4	Best K 10	5.205369	0	0.025321
3	All CP	1.276155	0.201901	0.033617
2	Proposal 10	0.957116	0.338509	0.041844
1	Best K 50	0.151124	0.879878	0.05

Table 2: Post Hoc comparison Table for  $\alpha = 0.05$  (FRIEDMAN)

Finner's procedure rejects those hypotheses that have an unadjusted p-value  $\leq 0.033617$ .

### 3 Adjusted P-Values (Friedman)

Adjusted P-values obtained through the application of the post hoc methods (Friedman).

i	algorithm	unadjusted $p$
1	BestCP	0
2	Covering	0
3	Best K 10	0
4	All CP	0.201901
5	Proposal 10	0.338509
6	Best K 50	0.879878

Table 3: Adjusted  $p$ -values (FRIEDMAN) (I)

i	algorithm	unadjusted $p$	$p_{Finner}$
1	BestCP	0	0
2	Covering	0	0
3	Best K 10	0	0
4	All CP	0.201901	0.287007
5	Proposal 10	0.338509	0.390984
6	Best K 50	0.879878	0.879878

Table 4: Adjusted  $p$ -values (FRIEDMAN) (II)