# Evolutive Algorithms Applied To The Straight Line Segment Classifier

Rosario A. Medina Rodríguez and Ronaldo Fumio Hashimoto
rmedinar@vision.ime.usp.br, ronaldo@ime.usp.br

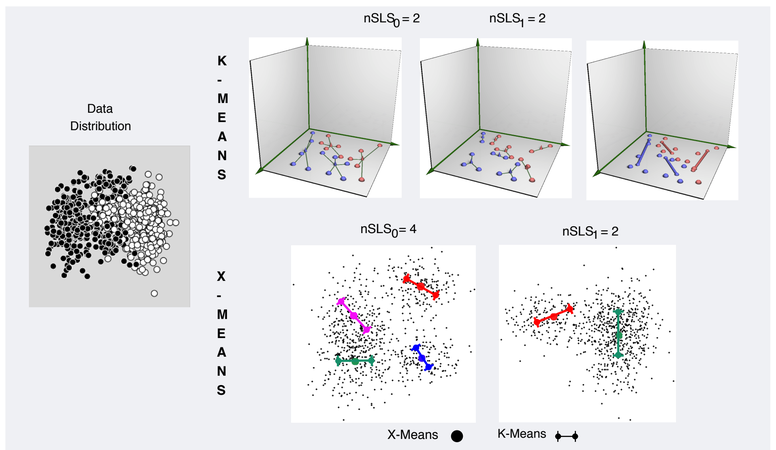
FAPESP

## Contributions

*We introduce a new training algorithm for the SLS binary classifier. A combination of evolutive algorithms and Gradient Descent method is used to improve its accuracy. In addition, we estimate the number of straight line segments by applying the clustering algorithm X-Means.*
*Our approach showed improvements in the accuracy compared to the original algorithm when applied to synthetic and real datasets. Also, our results are comparable with the SVM ones.*

## Selecting the Number of SLSs

| SLSs | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | **1-1** | 1-2 | 1-3 | 1-4 |
| 2 | 2-1 | **2-2** | 2-3 | 2-4 |
| 3 | 3-1 | 3-2 | **3-3** | 3-4 |
| 4 | 4-1 | 4-2 | 4-3 | **4-4** |




## Hybrid Optimization Methods

Evolutive algorithms can assist the Gradient Descent method by providing a new set of initial positions for $\mathcal{L}_0$ and $\mathcal{L}_1$, avoiding a local minima.

**Methods**:
**1.** Gradient Descent - GD
**2.** Genetic Algorithms - GA
**3.** Genetic Algorithms and Grad. Desc. - GA+GD
**4.** Dialectical Opt. and Grad. Desc. - DOM+GD
**5.** K-Beam Search and Grad. Desc. - KBM+GD

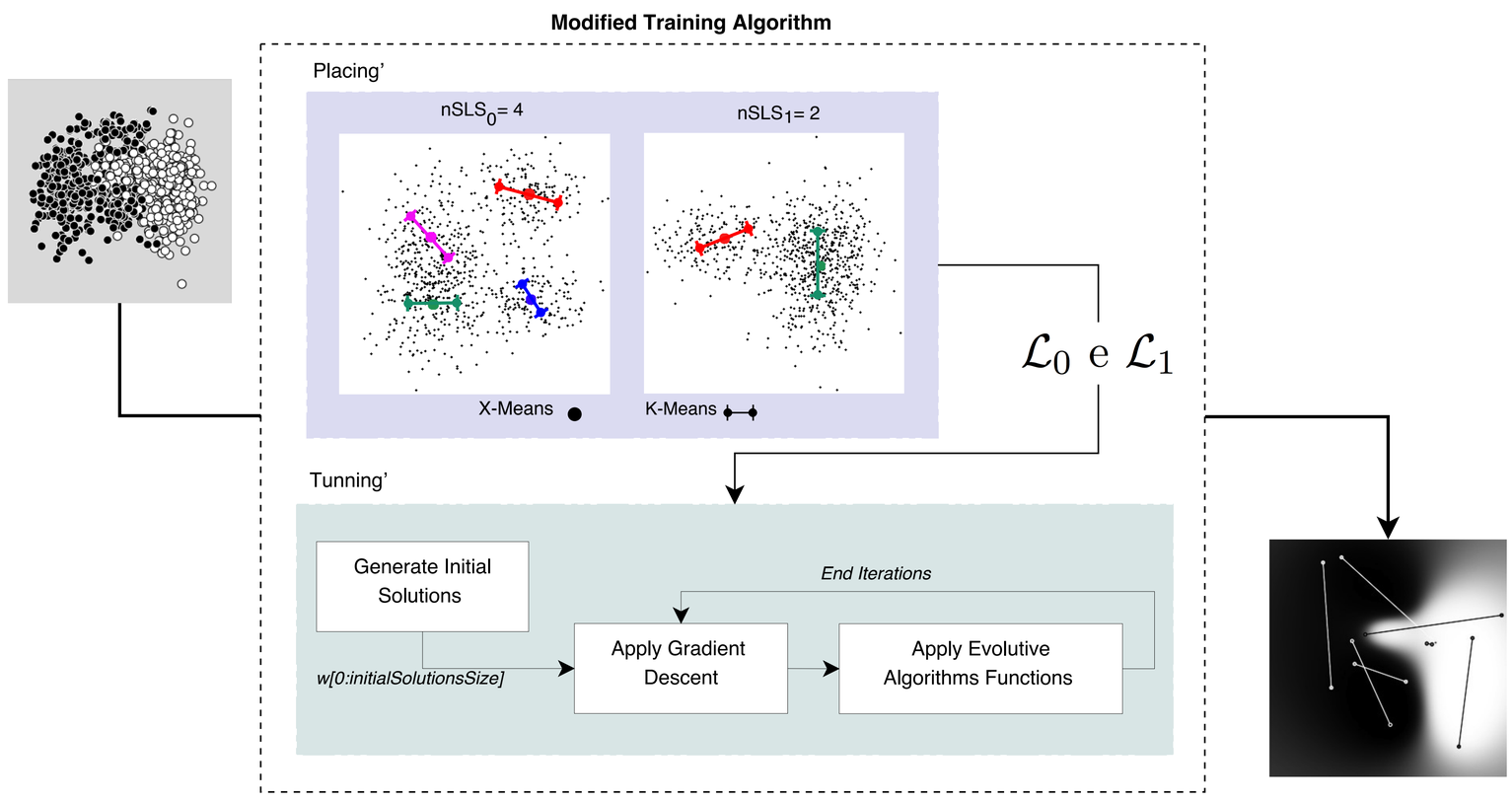
**Setup**:
**1.**Chromosome/Poles/States: Initial solutions, consisting of the concatenated extremities from $\mathcal{L}_0$ and $\mathcal{L}_1$ (i.e. $\left[\mathcal{L}_0|\mathcal{L}_1\right]$).
**2.**Fitness Function/Social Force/Cost Function: Represented by the MSE function:
$E(F_{\mathcal{L}_0,\mathcal{L}_1}) = \frac{1}{n}\sum_{i=1}^{n}[F_{\mathcal{L}_0,\mathcal{L}_1}(x_i) - y_i]^2$,
which is the classification function from the SLS classifier.

**Algorithm:**
**Require:** iterations, initialSolutionSize, minimumError
**Ensure:** $minValue$; optimized value.
1: w[0:initialSolutionSize] ← generateInitialSolutions()
2: **while** (it < iterations) **do**
3: **if** ($minValue$ > minimumError) **then**
4: w ← applyGradientDescent(w)
5: $minValue$ ← evolutiveAlgorithmFunctions(w)
6: **end if**
7: **end while**
8: **return** $minValue$

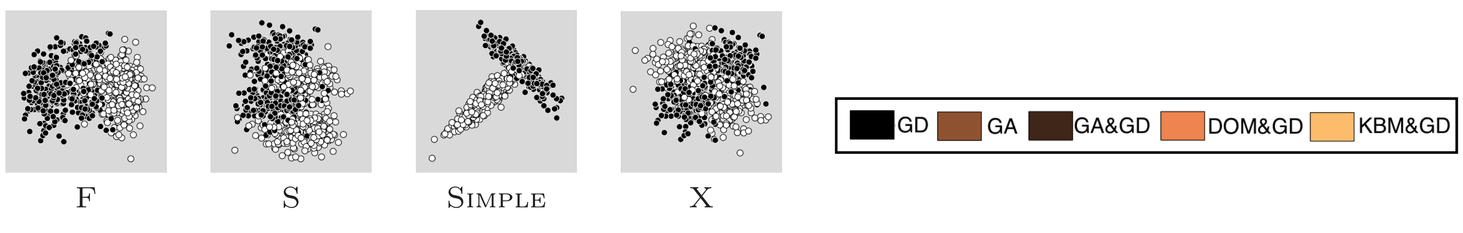
## References

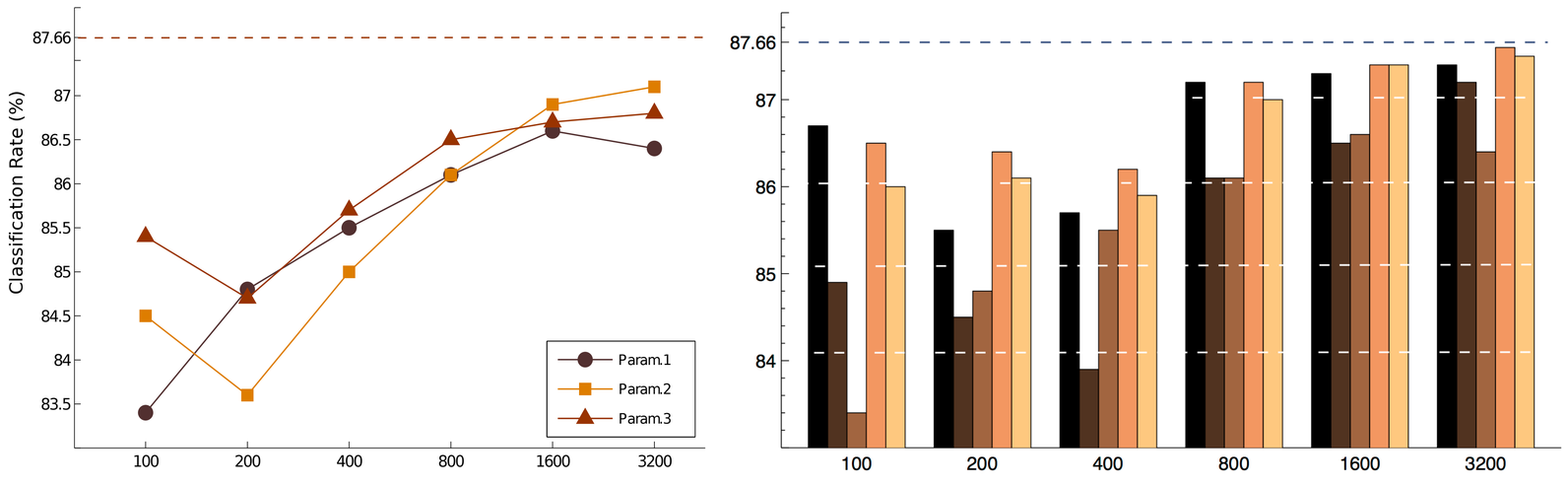
[1] D. Pelleg and A. Moore, "X-means: Extending k-means with efficient estimation of the number of clusters" in Proceeding of the 17th International Conference on Machine Learning. 2000.
[2] J. Ribeiro and R. Hashimoto, "Pattern Recognition Based on Straight Line Segments", Pattern Recognition, Recent Advances. 2010.
[3] W. D. Santos and F. D. Assis, "Optimization based on dialectics" in Proceedings of the 2009 International Joint Conference on Neural Networks. 2009.

## Methodology




## Results: Synthetic Distributions




## Results: Public Distributions

| Data Sets | SLSs - Number | |
|---|---|---|
| | X-Means | K-Means |
| australian | 9-1 | 8-8 |
| breast-cancer | 3-1 | 4-4 |
| diabetes | **3-5** | 1-1 |
| german | **5-3** | 2-2 |
| heart | 3-2 | 10-10 |
| ionosphere | 2-8 | 10-10 |
| liver-disorders | **4-2** | 3-3 |
| sonar | **2-2** | 4-4 |

| Public Data Sets | Proposal - Accuracy (%) | | | | | Best % - Original | |
|---|---|---|---|---|---|---|---|
| | GD | AG+GD | AG | MDO+GD | KBM+GD | Class. SLS | SVM |
| australian | **76.9** (0.6) | 75,0(2,7e-04) | 75,0 (2,7e-04) | 75,0 (1,9e-04) | 75,4 (0,6) | 87.0 (1.8) | 87.4 (1.6) |
| breast-cancer | **76.9** (2.4) | 75.2 (0.4) | 75.0 (0.0) | 75.0 (0.0) | 75.0 (0.0) | 98.1 (0.7) | 97.9 (0.9) |
| diabetes | 75.9 (0.2) | 81.2 (0.9) | 75.1 (0.03) | **81.5** (1.1) | 81.4 (0.7) | 76.4 (1.8) | 77.8 (1.8) |
| german | 76.5 (0.5) | 80.5 (0.8) | 75.2 (0.03) | 80.7 (1.1) | **80.9** (0.7) | 76.7 (2.2) | 77.3 (0.5) |
| heart | 75.5 (0.5) | 77.0 (1.1) | 75.0 (0.01) | **77.9** (1.7) | 77.1 (0.9) | 82.2 (3.3) | 85.1 (3.3) |
| ionosphere | 81.0 (3.5) | 94.7 (1.1) | 74.5 (3.7) | **95.1** (1.1) | 93.1 (1.8) | 95.2 (2.6) | 96.0 (2.1) |
| liver-disorders | 75.8 (0.5) | **78.9** (1.9) | 75.5 (0.3) | 77.3 (1.8) | 76.3 (1.2) | 70.1 (2.8) | 72.7 (2.7) |
| sonar | 80.9 (1.8) | **88.9** (0.8) | 75.3 (0.09) | 87.6 (2.8) | 87.8 (1.4) | 86.3 (4.1) | 88.4 (4.2) |

## Conclusions

*The application of evolutive optimization algorithms combined with the gradient descent method allows the accuracy improvement of the SLS classifier. We also suggest the X-Means algorithm to estimate the number of SLSs.*