Large margin classification with graph-based models

Prof. Antônio de Pádua Braga
Universidade Federal de Minas Gerais - UFMG
http://www.ppgee.ufmg.br/~apbraga

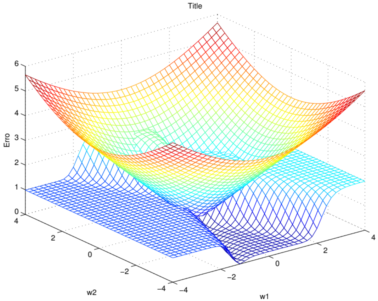
Federal University of Minas Gerais, Brazil

Outline

Introduction

Supervised learning

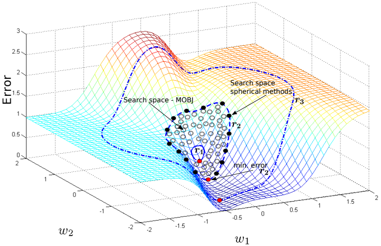
- Optimization problem;
- Multi-objective:
  - Trade-off objective functions.



Introduction

Supervised learning

- Feasible solutions;
  - Equality or inequality constraints;
  - Requires prior or posterior selection criteria.



Multi-objective learning

- HP Rocha, MA Costa, AP Braga (2020) Neural networks multiobjective learning with spherical representation of weights (2020). IEEE transactions on neural networks and learning systems 31 (11), 4761-4775.
- . . .
- R de Albuquerque Teixeira, AP Braga, RHC Takahashi, RR Saldanha (2000). Improving generalization of MLPs with multi-objective optimization, Neurocomputing 35 (1-4), 189-194

Introduction

SVM's dual formulation

Given vector of parameters p and C, maximize:

$$W(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j K(\mathrm{x}_i, \mathrm{x}_j, \mathrm{p})$$

subject to

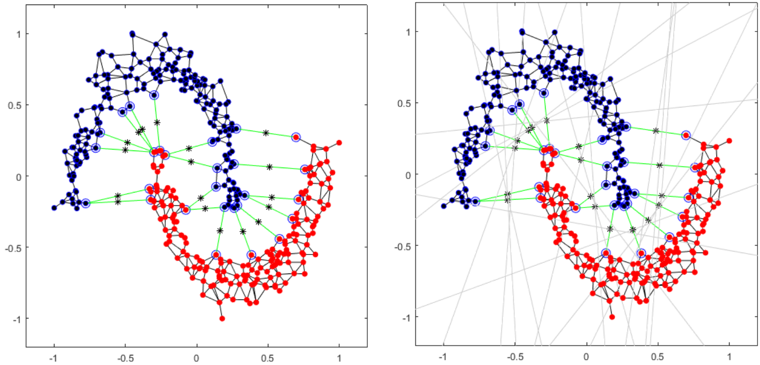
- $\sum_{i=1}^{N} \alpha_i y_i = 0$
- $0 \leq \alpha_i \leq C$

Prior parameters

- Kernel parameters p;
- Regularization parameter C.

Introduction

A graph-based model



Introduction

Graph-based methods

- Distance matrix calculation;
- Implementation of graph;

Implementation issues

- Parallelization;
- Convolution;
- IC implementation;
- Less parameters to set.

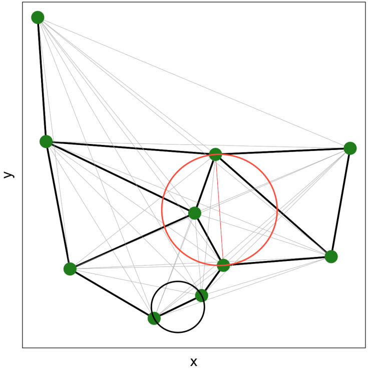
Aim

Small scale embedded learning systems.

Introduction

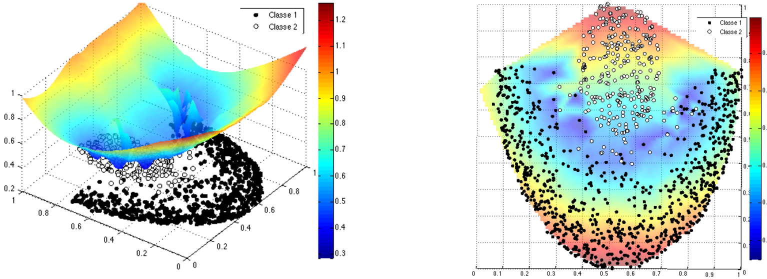
Gabriel Graph

- Planar graph
  - Locality/neighborhood information is preserved.
- Graph properties $\rightarrow$ properties of the data;
- Preserves the structure of the dataset.



Introduction

Properties of the graph $\rightarrow$ structural information of the data



Distance between classes

Mean distance of a sample to all others of the opposite class, considering graph paths.

Introduction

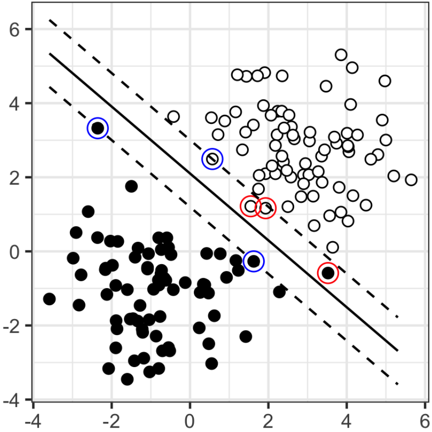
In this presentation

- Methods developed so far;
- Implementation issues;
- Discussions on further developments.

Outline

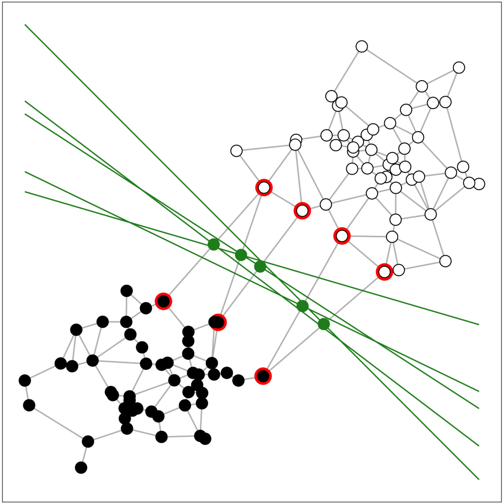
SVMs formulation

Large margin classification with SVMs



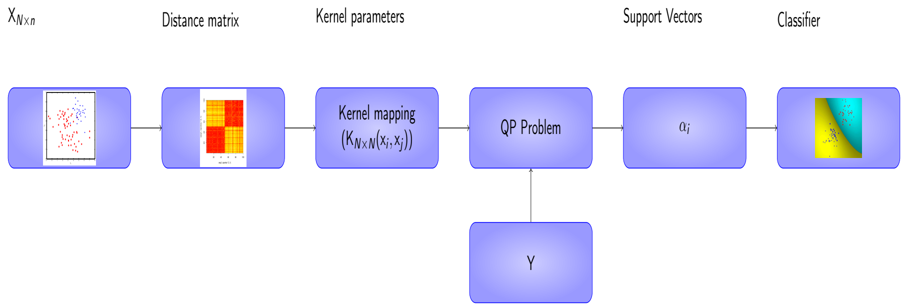
Graph-based approach

Large margin classification with a Gabriel graph



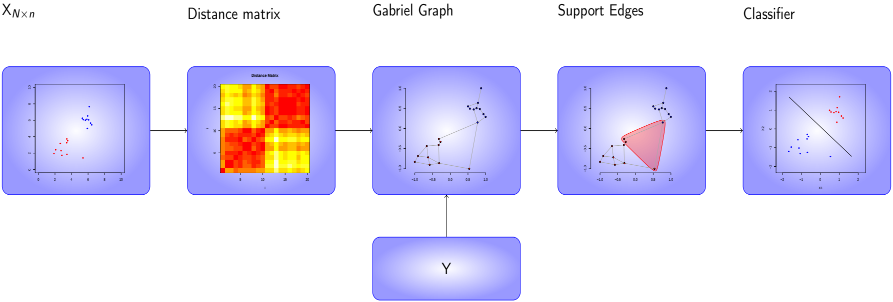
Kernel Mapping and SVM Learning

Feature space optmization



Gabriel Graph Learning

Input space classifier

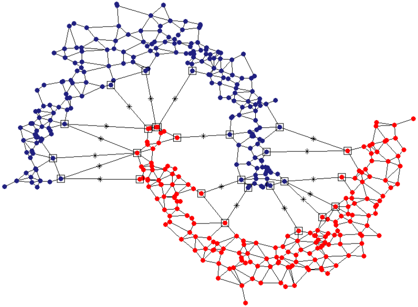


Outline

Chip Class

Principles

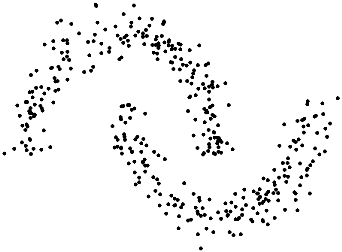
- Graph $\rightarrow$ structure of the data;
- Obtain SSVs and SEs;
- Bisecting hyperplanes equations for each SE are known;
- Combine hyperplanes or choose the nearest one.

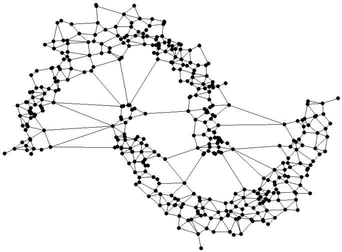


LCB Torres, CL Castro, F Coelho, FS Torres, AP Braga (2015). Distance-based large margin classifier suitable for integrated circuit implementation, Electronics Letters 51 (24), 1967-1969.

Gabriel Graph

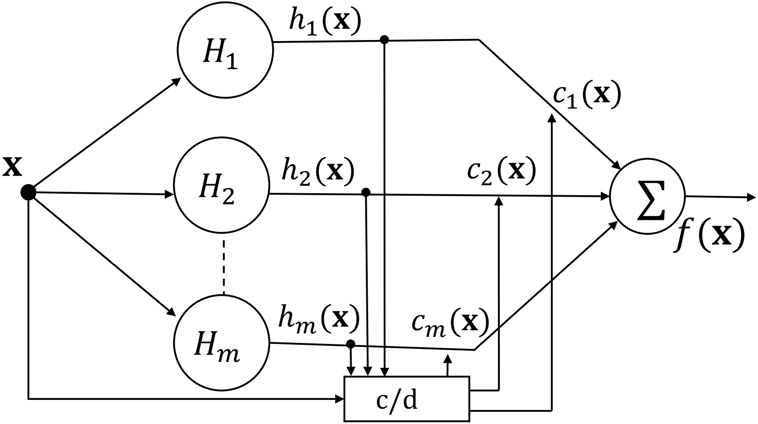
$$(\mathrm{v}_i, \mathrm{v}_j) \in \mathcal{E} \leftrightarrow \delta^{\mathbf{2}}(\mathrm{v}_i, \mathrm{v}_j) \leq \left[\delta^{\mathbf{2}}(\mathrm{v}_i, z) + \delta^{\mathbf{2}}(\mathrm{v}_j, z)\right] \ \forall\, z \in \mathcal{V}, \mathrm{v}_i, \mathrm{v}_j \neq z$$





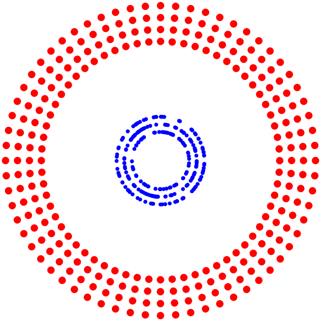
Chip Class

Combination of hyperplanes

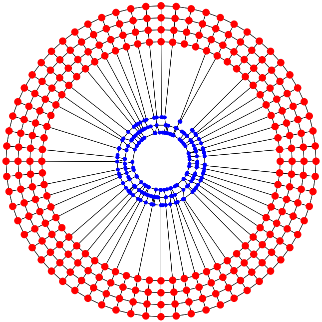


Chip Class

Dataset

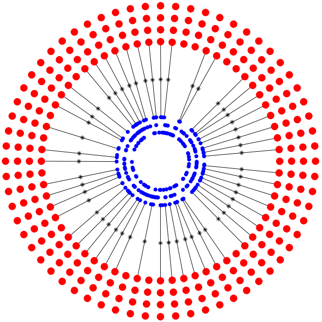


Gabriel Graph

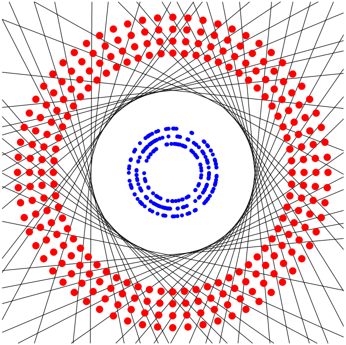


Chip Class

Support Edges

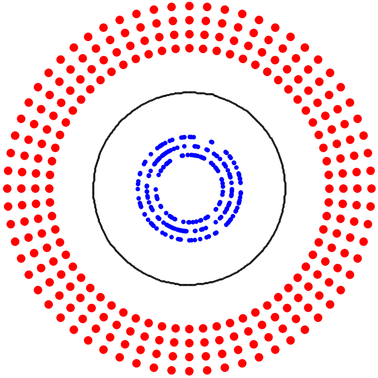


Hyperplanes



Chip Class

Combination of hyperplanes



Chip Class

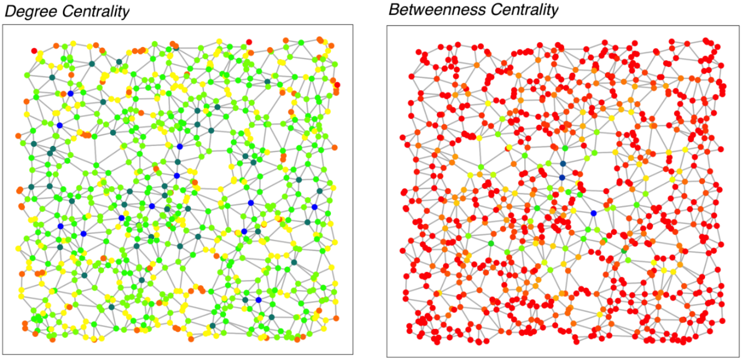
Example

Graph-based large margin classification.
Example of graph formation.

Outline

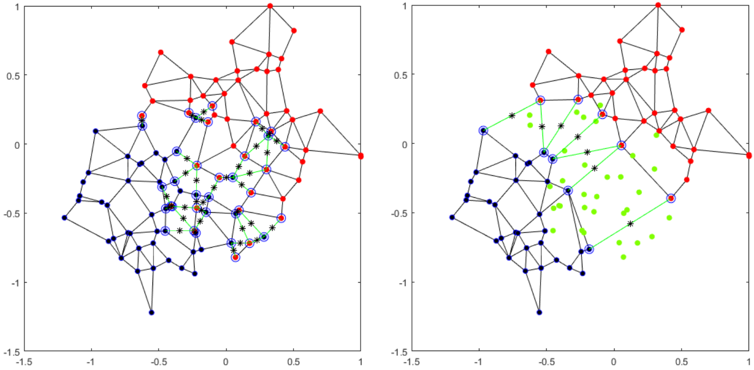
What can we gain with the graph representation?

Properties of the graph → structural information of the data



What can we gain with the graph representation?

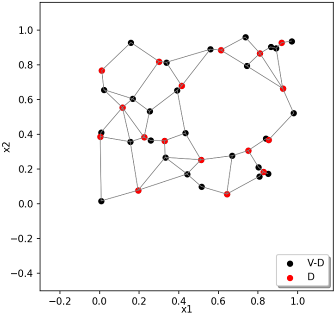
Regularization according the degree of a vertex

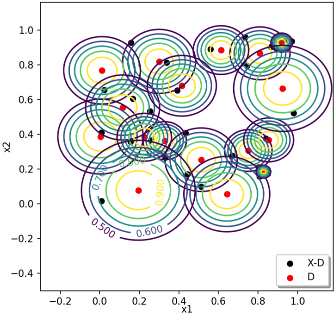


What can we gain with the graph representation?

Properties of the graph → structural information of the data

Dominating set → dataset coverage.



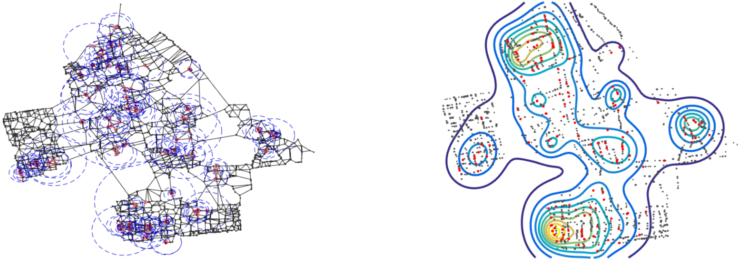


Model parameters are extracted directly from the adjacency matrix.

Outline

Dengue fever transmission

Six months data (Lassance, MG, Brazil)



Dengue fever transmission

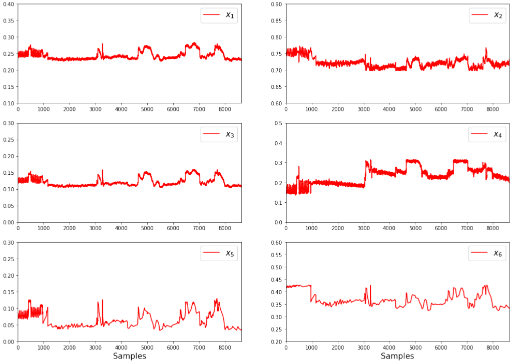
Example

Density estimation: example of geospatial temporal forecasting.
Video: Density Estimation.
Video: Contour plots.

Example of a real industrial application

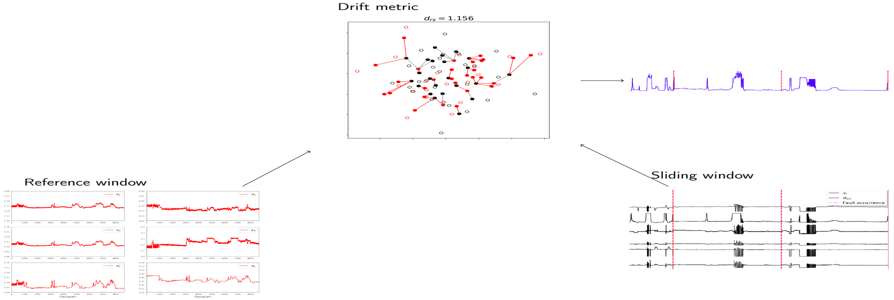
Drifts and anomaly detection of pumping motors of a large mining plant.

- Six variables of 8 motors: current, speed, torque, tank level, flow rate and fluid density;
- 145 days of data with 10s sampling rate.



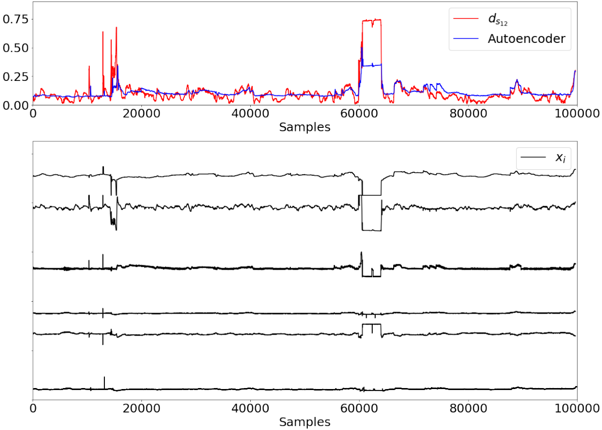
Example of a real industrial application

Drifts in dominating set → drifts in data



Example of a real industrial application

Drifts in dominating set → drifts in data



Alvarenga, W., Campos, F., Costa, A., Salis, T., Magalhães, E., Torres, L, Braga, A.P. (2021) Time Domain Graph-Based Anomaly Detection Approach Applied to a Real Industrial Problem (Under Review).

Example of a real industrial application

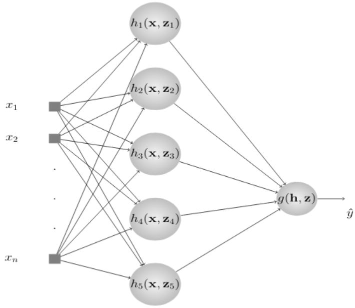
Discussions

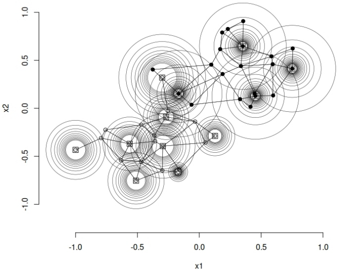
- Window size is the only parameter to set;
- Online/incremental learning: no need to update the whole graph;
- Performance/behaviour close to the autoencoder in other applications too;
- Computational cost: incremental implementation of DS;

Outline

RBF with centers located in DS

- Linear combination of RBFs located in DS.

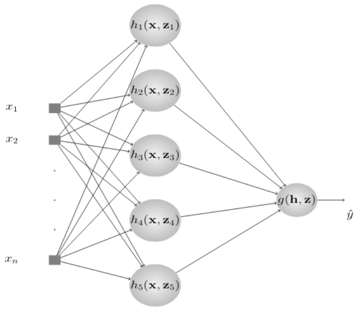


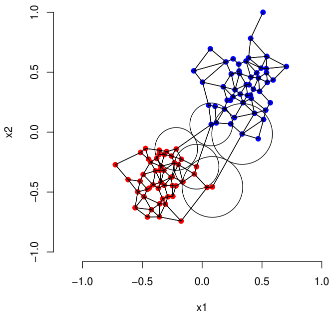


Queiroz, M., Coelho, F., Torres, L., Campos, F. Lara, G. Alvarenga, W. and Braga, A.P. (2021+). RBF neural networks design with graph based structural information from dominating sets. Under review.

Margin-based RBF

- Linear combination of RBFs located on SSV.

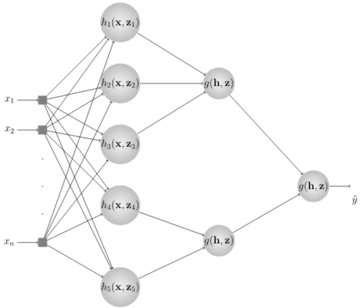


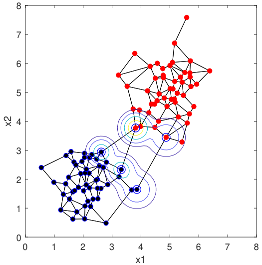


Luiz C. B. Torres, André P. Lemos, Cristiano Leite Castro, Antônio P. Braga (2014) A Geometrical Approach for Parameter Selection of Radial Basis Functions Networks. ICANN 2014: 531-538

Gaussian mixture of RBF functions

- Each SSV defines a multivariate normal distribution (MDF), which are combined for each class.

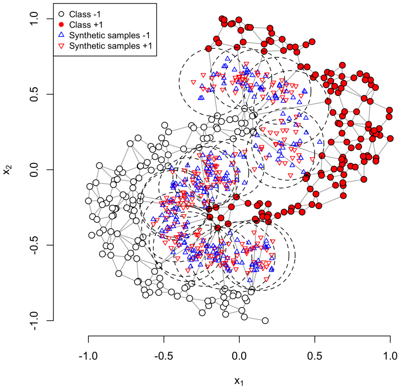




LCB Torres, CL Castro, F Coelho, AP Braga (2020). Large margin Gaussian mixture classifier with a gabriel graph geometric representation of data set structure. IEEE transactions on neural networks and learning systems 32 (3), 1400-1406

Data augmentation in the margin

Regularization, training with noise, data augmentation



AD Assis, LCB Torres, LRG Araújo, VM Hanriot, AP Braga (2021). Neural Networks Regularization With Graph-Based Local Resampling. IEEE Access 9, 50727-50737

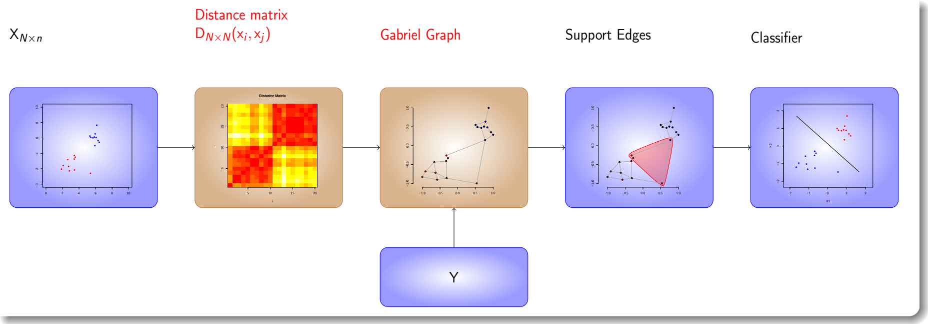
Overview of performance

AUC values.

Data Set	Chip Class	SVM-RBF	SVM-Poly	N/N_d
Australian Cr.	0,85±0,04	0,86±0,04	**0,87±0,04**	690/14
Banknote Auth.	0,98±0,03	**1±0**	**1±0**	1372/4
BcrHess	**0,81±0,12**	0,76±0,11	0,77±0,15	133/30
B. Cancer W.P	0,96±0,03	**0,97±0,01**	0,96±0,03	683/9
Climate M.S.C.	**0,84±0,07**	0,53±0,06	0,72±0,11	540/18
Fertility	**0,59±0,26**	0,5±0	0,5±0	100/9
German Cr	0,67±0,04	0,66±0,07	**0,68±0,05**	1000/24
Golub	0,77±0,17	**0,8±0,16**	0,78±0,17	72/50
Habermans S.	**0,57±0,09**	0,52±0,06	0,5±0,02	306/3
ILPD	**0,56±0,09**	0,49±0,02	0,5±0	579/10
Liver Disorders	0,61±0,1	0,67±0,05	**0,72±0,07**	345/6
P. ind. Diabetes	**0,72±0,04**	0,71±0,05	0,71±0,07	768/8
Parkinsons	**0,9±0,15**	0,77±0,11	0,81±0,12	195/22
Sonar, M vs. R.	**0,88±0,08**	0,84±0,09	0,87±0,08	208/60
Stalog Heart	0,8±0,08	**0,83±0,07**	**0,83±0,07**	270/13
Average Rank	1.87	2.23	1.90	

Outline

Bottlenecks



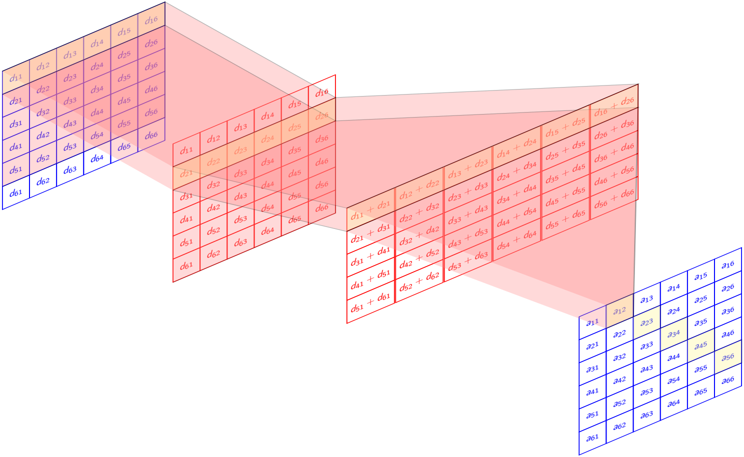
Implementation

- Distance matrix calculation can be parallelized
  - CUDA/GPU implementation.

Implementation

Convolutional implementation of the graph

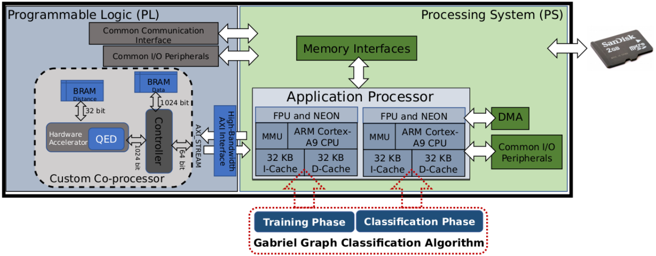
Gain is about 40 times compared to our best C++ implementation



Implementation

Hardware FPGA/Soc version of Chip Class

- Implementation on a Zynq-7000 SoC architecture



J Arias-Garcia, A Mafra, L Gade, F Coelho, C Castro, L Torres, A Braga (2020) Enhancing performance of gabriel graph-based classifiers by a hardware co-processor for embedded system applications. IEEE Transactions on Industrial Informatics 17 (2), 1186-1196

J Arias-Garcia, A Mafra, L Gade, J. Yudi, F Coelho, C Castro, L Torres, A Braga (2021+) Improved Design for Hardware Implementation of Graph-Based Large Margin Classifiers for Embedded Edge Computing". Under review.

R-package available

GGClassification available from CRAN repository

- **GG Fast**: Gabriel Graph generation function implemented in C++ and integrated to the main package with Rcpp;
- **DS**: returns the Dominating Set of a graph, given its adjacency matrix;
- **Chip Class**: Combination of hyperplanes classifier;
- **RBF Class**: margin-based RBF classifier;
- **RBFDS Class**: dominating set RBF classifier;
- **GMM Class**: Gaussian mixture RBF classifier.

Package webpage: https://cran.r-project.org/package=GGClassification

Outline

Discussions

- Representation of the problem, based on the assumptions:
  - Graph represents the problem;
  - Graph properties can be used to build the classifier;
- Graph is unique for a given dataset
  - Classifier is unique for a given distance metric and regularization threshold;
  - Little/no user interaction needed;
- Incremental/online implementation
  - No need to calculate the whole graph;

Discussions

- Offline implementation
  - Parallelization and convolution;
- In-circuit implementation
  - Distance calculations and convolutions;
  - Small scale embedded systems;
  - Our implementation on a SoC architecture: 1000 samples.

Acknowledgements

Thanks to my students and collaborators