Projetando rede neurais RBF através de informações estruturais dos dados e cobertura de conjuntos

Marcelo D. Queiroz* Luiz C. B. Torres** Frederico Coelho*
Antonio P. Braga*

* Universidade Federal de Minas Gerais
Departamento de Engenharia Eletrônica,
Av. Antônio Carlos 6627, 31270-901, Belo Horizonte, MG, Brasil
(e-mail: mdoqueiroz@gmail.com / fredgfc@ufmg.br / apbraga@ufmg.br)
** Universidade Federal de Ouro Preto
Departamento de Computação e Sistemas
João Monlevade - MG, Brasil (e-mail: luiz.torres@ufop.edu.br)

Abstract: The use of RBF (*Radial Basis Function*) in classification problems of samples has been a very discussed approach in the literature. Its diverse applications allow you to deal with problems of different complexities. One of the challenges in using this type of function is to find the appropriate parameters to avoid sub or oversize in the solution of the problem and in this context to define radius normally given by the constant sigma is another issue. This work proposes a methodology based on well-known mathematical postulates to find the sigma, for a given set of samples, generated in an autonomous way after using the Gaussians in the samples in a similar way to the other known methods.

Keywords: Gaussian, Gabriel's Graph, Dominant Set and Neural Network.

1. INTRODUÇÃO

Funções de base radiais (RBF) são largamente utilizadas na literatura para problemas de classificação e de reconhecimento de padrões, sobretudo quando as classes não são linearmente separáveis. Segundo o teorema de Cover (Haykin, 2009), um problema não linearmente separável se torna linearmente separável quando mapeado em uma dimensão suficientemente maior. Nas RBFs quem faz este mapeamento em dimensões mais elevadas é o kernel. Estas funções são conhecidas pela simplicidade de sua formulação matemática e sua habilidade em lidar com problemas multidimensionais de diferentes complexidades de solução. Pelo teorema de Micchelli, das diferentes formulações de RBFs, as mais estudadas baseiam-se em três tipos diferentes de funções (Haykin, 2005): Multiquadrática, inversa da Multiquadrática e Gaussiana. Esta última pode ser formulada como:

$$\varphi(x) = \exp\left(-\frac{x^2}{2\sigma^2}\right), \text{ para algum } \sigma > 0 \text{ e } x \in \mathbb{R} \quad (1)$$

Onde $\varphi(x)$ é a formulação da Gaussiana, x é variável aleatória de posicionamento ou o centro da Gaussiana e σ é a variância dentre suas possíveis formulações ou em outras palavras o raio de abertura da Gaussiana. Embora as RBFs de kernel Gaussiano sejam uma das mais eficientes metodologias de classificação de padrões, o seu uso implica nos ajustes de seus parâmetros x e σ para evitar sub ou superdimensionamento na solução do problema. Como pode ser visto em (1) a Gaussiana depende apenas destes

dois parâmetros, ou seja, penas as escolhas adequadas poderão permitir que a sua aplicação na classificação de padrões seja eficaz. A utilização dos Kernels para problemas complexos parte do princípio que os dados podem ser separados através de distribuições estatísticas em que as amostras deste subconjunto possuem verossimilhanças próximas, ou seja, que as amostras próximas tendem a pertencer ao mesmo conjunto. O uso de kernels gaussianos vem sendo aplicado em redes neurais com autoaprendizagem na solução de problemas de classificação sendo um dos campos abordados o ajuste dos parâmetros x e σ de (1). Há estudos recentes Torres (2012); Souza et al. (2014); Torres (2016); Wanderley (2013) com propostas para encontrar estes parâmetros de (1) sem a necessidade de arbitragem dos centros e com a arbitragem dos raios. Este trabalho pretende apresentar uma nova abordagem de cálculo destes parâmetros através dos postulados matemáticos conhecidos, grafo de Gabriel e teoria dos conjuntos dominantes, para encontrar-los de maneira dinâmica, ou sem arbitragem, em cada problema ajustando os kernels para cada solução e aplica-los em uma rede neural em problemas de classificação. Neste sentido o artigo está separado da seguinte forma: Seção 2. apresentamos a Metodologia proposta, na Seção 3. descrevemos a Preparação para os testes, na Seção 4. apresentamos os Resultados encontrados em três diferentes formas: Primeiro comparando outros métodos de kernel que arbitram os centros com as metodologias: *Affinity* Frey and Dueck (2007), *Silhouette* Rousseeuw (1987) e Aleatório; Segundo comparando com outros métodos de classificação utilizando base sintética; Terceiro comparando com outros métodos de classificação

clássicos: RBF, Bayesiano, MLP (*Multi Layer Perceptron*), ELM (*Extream Learn Machine*) e SVM (*Support Vector Machine*), utilizando três bases reais e por fim, na Seção 5. apresentamos nossas Conclusões.

2. METODOLOGIA

Dado um problema de classificação de N amostras com $\hat{y}$ rótulos conhecidos, a rede neural modelada deverá ser capaz de classifica-las sem necessidade de ajustar os parâmetros dos kernels radiais encontrados. Conforme Haykin (2005), o Kernel é uma forma de aprendizagem não supervisionada em que um conjunto de observações (ou seja, amostras) é dividido em agrupamentos naturais ou conjuntos de padrões de tal forma que a medida de semelhança entre quaisquer pares de observações atribuídas a cada conjunto minimize a função de custo Haykin (2009):

$$J(C) = \frac{1}{2} \sum_{j=1}^K \sum_{C(i)=j} \sum_{C(i')=j} \delta(x_i, x_{i'}) \quad (2)$$

onde K é o Kernel, $J(C)$ a função de custo, i e j os números das amostras, x_i e $x_{i'}$ os vetores de amostras, sendo δ a distância euclidiana medida entre os centros de cada kernel e as médias entre as amostras mais próximas, tal que quanto menor a distância, mais próxima a amostra estará do seu conjunto. O desafio então passa a ser: Como encontrar estes conjuntos em uma amostragem genericamente distribuída. Uma vez encontrados estes conjuntos poderemos então inferir os parâmetros das RBFs. Nossa proposta é utilizar o grafo de Gabriel e a teoria dos conjuntos dominantes para encontrar estes conjuntos.

A. Grafo de Gabriel

O Grafo de Gabriel $\ddot{G}$ é interpretado como um conjunto de pontos de um hiperespaço $\mathcal{H}$ formado em um espaço euclidiano representado pelo plano $\mathcal{S}$ expresso pela medida da distância média entre estes pontos. Conforme Zhang and King (2002), considerando um grafo com $\mathcal{P} = \{x_1, \dots, x_n\}$ pontos e dois destes, x_i e x_j , conectados por uma aresta α , o grafo $\ddot{G}$ existirá apenas se ao se traçar uma esfera τ de diâmetro (x_i, x_j) não existir nenhum outro ponto de $\mathcal{P}$ interno à esfera. O Grafo de Gabriel $\ddot{G}$ pode ser descrito como um sub-conjunto de pontos do Diagrama de Voroni $V(\mathcal{S})$ e também um sub-grafo da triangulação de Delaunay $TD(\mathcal{S})$ Souza et al. (2014); Zhang and King (2002). Conforme Berg et al. (2008) O Grafo de Gabriel $\ddot{G}(\mathcal{S})$ de um conjunto de pontos de um plano $\ddot{G}(v, \alpha) \in \mathcal{S}$, onde v são os pontos que representam o conjunto dos vértices V do grafo e α suas arestas, deve obedecer à seguinte definição Torres (2012, 2016):

$$(v_i, v_j) \in \alpha \leftrightarrow \delta^2(v_i, v_j) \leq [\delta^2(v_i, x) + \delta^2(v_j, x)] \quad \forall x \in V, v_i, v_j \neq x \quad (3)$$

A Figura 1 demonstra o processo de formação do grafo de Gabriel. Em 1(i) há uma distribuição genérica de pontos $\mathcal{P} = \{x_1, \dots, x_n\}$ do plano $\mathcal{S}$. Considera-se cada um dos pontos como vértice de um grafo não orientado. Calcula-se a distância euclidiana entre os vértices. Interliga-se o

par de vértices com uma hipersfera τ (círculo tracejado) de diâmetro dado pela distância euclidiana entre cada par, conforme 1(ii). Caso não exista nenhum outro vértice $x \in \mathcal{P}$ dentro da hipersfera é traçada uma aresta α interligando estes vértices, formando $\ddot{G}_{(\mathcal{S})}(v, \alpha)$ demonstrado em 1(iv) e 1(v). Na formação do grafo, pontos internos ou externos, ou que não estão circunscritos na hipersfera como demonstrado em nas duas esferas de 1(ii), podem não satisfazer (3). Neste caso por $TD(\mathcal{S})$ não haverá uma aresta entre os vértices, sendo feito o processo traçando novas hipersferas de raio menor como demonstrado em 1(iii) e novas arestas como mostrado em 1(iv).

Figura 1. Formação do Grafo de Gabriel

B. Teoria do Conjunto Dominante

Um conjunto dominante em um grafo $\ddot{G}_{(\mathcal{S})}(v, \alpha)$ é um conjunto de vértices V do plano $\mathcal{S}$ tal que $\mathcal{S} \subseteq V(\ddot{G}_{(\mathcal{S})})$, de tal maneira que todo vértice do grafo ou pertence a $\mathcal{S}$ ou é adjacente a pelo menos um elemento de $\mathcal{S}$ Sousa (2015). Conjuntos dominantes estão intimamente relacionados aos conjuntos independentes. Um conjunto independente é aquele que possui dois vértices de $G_{(\mathcal{S})}$ em $\mathcal{S}$ tal que não exista, dois vértices adjacentes contidos em $\mathcal{S}$. Um conjunto independente é também um conjunto dominante se, e somente se, é um conjunto independente máximo, de modo que qualquer conjunto independente máximo em um grafo é necessariamente também um conjunto dominante mínimo. Assim, o menor conjunto independente máximo também é o menor conjunto dominante independente. Em outras palavras, dado o grafo $\ddot{G}_{(\mathcal{S})}(v, \alpha)$, este será um subconjunto H de $\ddot{G}_{(\mathcal{S})}$ de tal modo que cada vértice v que não está em H é adjacente a pelo menos um membro de $\ddot{G}_{(\mathcal{S})}$. Um grafo pode variar o número de dominação definido pelas suas interligações, sendo γ o número mínimo e Γ o número máximo. A Figura 2 mostra um grafo

com diferentes dominações: $\gamma = \{1, 3, 5\}(ii)$, com $\gamma = \{3, 6, 7, 8\}(iii)$ e com $\gamma = \{2, 4, 6, 7, 8\}(iv)$, tal que para este grafo o dominante mínimo será $\gamma(\ddot{G}_{(S)}) = 3$ e o dominante máximo será $\Gamma(\ddot{G}_{(S)}) = 5$.

Figura 2. Conjunto Dominante mínimo e máximo

C. Método Proposto

Com os conceitos descritos nas seções A. e B., o método propõe utilizar os vértices dos subconjuntos dominantes encontrados a partir do grafo de Gabriel como centros c_i das Gaussianas da função radial e a média da distância euclidiana entre estes vértices dominantes e suas adjacências o raio σ_i . Assim, $\forall i = \{1, 2, \dots, n\}$ de x_1 e x_2 temos:

$$c_i(x_1) = \mathcal{V}_{D_o(i)}(x_1) \text{ e } c_i(x_2) = \mathcal{V}_{D_o(i)}(x_2) \quad (4)$$

$$\sigma_i = \frac{1}{m} \sum_{i=1}^m \|\mathcal{V}_{D_o(i)}(x) - \mathcal{V}_{Adj(i)}(x)\| \quad (5)$$

A Figura 3 ilustra o processo. Esta proposição permite encontrar para qualquer agrupamento os centros c e os raios σ das Gaussianas sem a necessidade de ajustes destes parâmetros descritos em (1).

Figura 3. Parâmetros da RBF

Para o método proposto foi elaborado um algoritmo que separa as amostras x_i e x_j em Kernels radiais através dos conceitos descritos nas seções A. e B. Primeiramente ele foi testado em uma base sintética e depois aplicado a bases conhecidas na literatura, seguindo os passos propostos no Algoritmo 1.

Algoritmo 1: RBF Proposto

```

1  carrega a base de dados
2  mistura dos dados da base
3  divide amostras (70% Treina e 30% teste)
4  para -Repete até o critério de parada faça
5      gg(x1,x2) - Monta o Grafo de Gabriel
6      para x1 = 1 a N faça
7          para x2 = 1 a N faça
8              evolui  $\ddot{G}(x1, x2)$  equação 3;
9          fim
10     fim
11     Do(gg1,gg2) - Vértices dominantes
12     para gg1 = 1 a N faça
13         para gg2 = 1 a N faça
14             evolui  $\gamma(gg1, gg2)$ ;
15         fim
16     fim
17 fim
18 Separa os Dominantes  $\mathcal{V}_{D_o(i)}(x)$  e encontra
    as adjacências  $\mathcal{V}_{Adj(i)}(x)$ ;
19 Calcula a dist. Euclidiana entre
    ( $\mathcal{V}_{D_o(i)}, \mathcal{V}_{Adj(i)}$ );
20 Encontra as RBF com os centros  $c = \gamma$  e
     $\sigma = \text{mean}(\text{dist}(\mathcal{V}_{D_o(i)}, \mathcal{V}_{Adj(i)}))$ ;
21 Inicia o treinamento da rede;
22 Encontra os pesos  $w$ ;
23 Inicia o teste com os  $w$  encontrados;
24 Retorna AUC de treinamento, AUC de
    teste e acurácia de teste.
```

A Figura 4 demonstra graficamente o resultado da seleção de centros (Gaussianas) para uma base sintética bidimensional genericamente distribuída, após a separação dos vértices dominantes do grafo de Gabriel.

Figura 4. Grafo de Gabriel e Conj. Dominantes

3. PREPARAÇÃO PARA OS TESTES

Para comparar os resultados obtidos pelo método proposto foram feitas duas abordagens distintas: A primeira, descrita na seção 4.2, faz uma comparação com outros métodos Gaussianos que encontram o número de centros c através dos modelos: *affinity* usando das amostras mais

próximas por afinidade sendo um processo recursivo, *silhouette* através de amostras verossimilhanças com kernels mais distintos ou definidos e aleatoriamente. Em conjunto com estes, foram usados os métodos *k-means* identificado com a letra *K* e *c-means*, identificado por *FCM*, para encontrar a localização dos centros *c*. Para arbitrar os raios foi utilizada a metodologia encontrada em Torres (2016), onde σ é igual à duas vezes o espaçamento médio entre os centros que equivale ao diâmetro das hiper-esferas traçadas entre cada vértice do grafo de Gabriel 1, ou seja: $\sigma_i = \phi_T$. A segunda, descrita na seção 4.3, faz uma comparação do resultado da classificação do método proposto com outros 5 métodos de classificação clássicos conhecidos na literatura: RBF-radial, *Bayesiano*-KDE, MLP (*Multi layer Perceptron*), ELM (*Extream learning Machine*) e SVM (*Support Vector Machine*). Todos os modelos utilizaram pacotes Caret, RSNNS e Klar no Rx64 versão 3.3.2. Foi possível perceber que cada um dos métodos testados necessitou de certo conhecimento para ajustar os parâmetros e adequar suas aplicações nos problemas de classificação tratados. Tais ajustes visam alcançar as melhores métricas, que neste caso são: AUC de treinamento, AUC de teste e acurácia de teste e para tal, cada pacote, sugere valores já testados e até pré configurados. Um ajuste adicional foi feito escolhendo funções RBF radiais para aproximar os modelos testados. Para os testes foram escolhidas uma base sintética onde será variado o número de amostras e três bases reais do repositório *UCI dataset- Machine Learning Repository* descritas no Quadro 1 abaixo:

Base de dados	Alias	N_d	N	N_1	N_2
Breast Cancer	<i>Breast Cancer</i>	9	683	444	239
Sonar Mines vs Rocks	<i>Sonar</i>	60	207	97	111
Parkinsons	<i>Parkinsons</i>	22	195	147	48

Quadro 1 - Bases reais testadas (*Repositório UCI*)

Onde N é o número total de amostras, N_d o número de dimensão da base e N_1 e N_2 o número de amostras de cada classe.

4. RESULTADOS

4.1 Ajustes para o treinamento e testes

Cada método foi testado 10 vezes com validação cruzada e os valores médios de cada resposta foram anotados nas tabelas. Para todos os testes foram separados 70% das amostras para treinamento da rede e 30% para o teste. Cada método foi ajustado conforme recomendado nos pacotes para garantir a a melhor acurácia dentro de suas proposições.

4.2 Primeira Abordagem

Com o objetivo de validar o modelo proposto com outros modelos Gaussianos conforme descrito na seção 3, foi feito um *benchmarking* usando as bases reais estudadas. A Tabela 1 demonstra os resultados encontrados.

O modelo proposto apresentou melhores AUC e acurácias de teste para as três bases o que validando a sua eficácia para bases de dimensões diferentes. Para a base *Breast*

Tabela 1. *Benchmarking* de métodos Gaussianos

Base	Modelo	AUC		Acurácia Teste
		Treina	Teste	
<i>Breast Cancer</i>	RBF-Proposto	0,9767	0,9814	0,9805
	K- <i>Silhouette</i>	0,9611	0,9597	0,9659
	K- <i>Affinity</i>	0,9706	0,9746	0,9756
	K-Aleatório	0,9969	0,8408	0,8732
	FCM- <i>Silhouette</i>	0,9614	0,9624	0,9707
	FCM- <i>Affinity</i>	0,9743	0,9488	0,9512
	FCM-Aleatório	0,9906	0,9288	0,9415
<i>Sonar</i>	RBF-Proposto	0,8096	0,8226	0,8226
	K- <i>Silhouette</i>	0,5522	0,5368	0,5645
	K- <i>Affinity</i>	0,8400	0,7419	0,7419
	K-Aleatório	0,9787	0,7594	0,7581
	FCM- <i>Silhouette</i>	0,5311	0,5479	0,5323
	FCM- <i>Affinity</i>	0,8306	0,7073	0,7097
	FCM-Aleatório	0,9451	0,7376	0,7419
<i>Parkinsons</i>	RBF-Proposto	0,9099	0,7920	0,8276
	K- <i>Silhouette</i>	0,7298	0,6375	0,7931
	K- <i>Affinity</i>	0,7951	0,6479	0,7931
	K-Aleatório	0,8658	0,7092	0,7586
	FCM- <i>Silhouette</i>	0,6761	0,7111	0,7759
	FCM- <i>Affinity</i>	0,7321	0,5987	0,7069
	FCM-Aleatório	0,8596	0,6651	0,7931

Cancer o pior resultado foi para o modelo K-Aleatório. Para a base *Sonar* o pior resultado foi para a abordagem FCM-*Silhouette* e para a base *Parkinsons* o pior resultado ficou com a abordagem FCM-*Affinity*. Estes resultados, demonstram que a depender do problema, cada método terá uma eficácia diferente, ou em outras palavras não há um modelo que seja o melhor para todos os problemas.

4.3 Segunda Abordagem

A segunda abordagem foi dividida em duas etapas, conforme descrito na seção 3. A primeira feita com uma base sintética dividida em duas classes normalmente distribuídas e a segunda com as bases reais descritas. Em ambas os resultados foram comparados usando os cinco modelos de classificação clássicos da literatura, sendo dois Gaussianos (RBF-Radial e KDE-*Bayesiano*), MLP, ELM e SVM, este último conhecido como uma das melhores abordagens de classificação da literatura.

A. Análise com Base sintética

Foi feita a variação do número de amostras de: $100 < n < 2000$, onde n é o número de amostras. Os resultados foram colocados nas tabelas 2 e 3. A tabela 2 nos mostra que os resultados foram muito próximos. Todos os métodos demonstraram eficácia sendo capazes de separarem as amostras com acurácias de teste próximas de 100%. O método proposto alcançou AUC e acurácias de teste muito próximas dos modelos existentes o que também pode ser observado na tabela 3, mas em ambos os testes fica sempre abaixo da abordagem SVM.

Isso nos permite dizer que o método proposto alcançou o objetivo de classificar as amostras com mínimo erro mesmo com a variação do número de amostras o que aumenta a complexidade do problema, sem qualquer necessidade de ser mexer em parâmetros. Neste sentido foi possível comprovar, ao se aplicar a metodologia proposta nestas

Tabela 2. Análise da base para $n=(100 \text{ e } 200)$

Modelo	n= 100			n= 200		
	AUC		Acurácia	AUC		Acurácia
	Treina	Teste		Treina	Teste	
RBF Proposto	0,9935	0,9944	0,9836	0,9981	0,9977	0,9667
RBF Radial	0,9887	0,9983	0,9688	0,9996	0,9975	0,9979
Bayesiano KDE	0,9878	0,9889	0,9750	0,9791	0,9965	0,9647
SVM	0,9942	0,9952	0,9995	0,9885	0,9999	0,9865
MLP	0,9945	0,9922	0,9899	0,9989	0,9983	0,9988
ELM	0,9984	0,9593	0,9615	0,9996	0,9779	0,9778

Tabela 3. Análise da base para $n=(1000 \text{ e } 2000)$

Modelo	n= 1000			n= 2000		
	AUC		Acurácia	AUC		Acurácia
	Treina	Teste		Treina	Teste	
RBF Proposto	0,9997	0,9927	0,9833	0,9993	0,9914	0,9717
RBF Radial	0,9984	0,9992	0,9766	0,9988	0,9965	0,9788
Bayesiano KDE	0,9844	0,9950	0,9802	0,9814	0,9991	0,9788
SVM	0,9961	0,9984	0,9999	0,9973	0,9988	0,9765
MLP	0,9996	0,9974	0,9995	0,9988	0,9988	0,9798
ELM	0,9988	0,9959	0,9748	0,9990	0,9989	0,9692

bases, que o uso do sub-grafo dominante do grafo de Gabriel para encontrar os centros e raios da Gaussianas na escolha das RBFs encontrou resultados próximos aos dos demais métodos cujos ajustes são determinantes para esta assertividade. A figura 5 mostra a superfície de separação traçada pelo método proposto com ($n = 100$) amostras, onde se observa que ocorreu um erro de classificação de pelo menos uma amostra de teste.

B. Análise com bases reais

Com o objetivo de validar a eficácia do método proposto em bases reais, foram feitos os testes com os cinco classificadores clássicos descritos em 3. A Tabela 4 mostra o *benchmarking* com os modelos de classificação para as três bases estudadas.

1. Base Breast Cancer

A tabela 4 mostra que para a base *Breast Cancer*, a abordagem proposta ficou com AUC e acurácia abaixo apenas do modelo SVM que é conhecido pela robustez na solução de problemas de classificação. Estes resultados demonstram que o modelo proposto conseguiu se aproximar dos resultados de modelos Gaussianos e classificadores clássicos conhecidos na literatura o que nos leva a dizer que alcançou o objetivo proposto.

Figura 5. Superfície de Separação

2. Base Sonar

Para a base Sonar, na Tabela 4, novamente pode ser observado que o modelo proposto ficou abaixo apenas da abordagem SVM. Chama a atenção o fato do modelo ter resultados proporcionais acima dos Gaussianos RBF-Radial e *bayesiano-KDE*, que apresentaram os piores resultados, quando se comparado com o MLP e ELM. Este fato demonstra uma vantagem da metodologia proposta entre os demais Gaussianos por sua eficácia nos resultados.

Tabela 4. Testes com as bases reais do UCI

Base	Modelo	AUC		Acurácia Teste
		Treina	Teste	
Breast Cancer	RBF-Proposto	0,9767	0,9684	0,9599
	RBF-Radial	0,9643	0,9568	0,9493
	Bayesiano KDE	0,9704	0,9507	0,9465
	SVM	0,9762	0,9701	0,9739
	MLP	0,9709	0,9523	0,9561
	ELM	0,9942	0,9670	0,9569
Sonar	RBF-Proposto	0,8096	0,8226	0,8226
	RBF-Radial	0,9157	0,7753	0,7742
	Bayesiano KDE	0,9155	0,7564	0,7521
	SVM	0,9919	0,8601	0,8710
	MLP	0,8490	0,8170	0,7903
	ELM	0,9783	0,7836	0,7919
Parkinsons	RBF-Proposto	0,9099	0,7920	0,8276
	RBF-Radial	0,7724	0,7901	0,7857
	Bayesiano KDE	0,7098	0,7488	0,7939
	SVM	0,9147	0,9510	0,8417
	MLP	0,8391	0,8697	0,7758
	ELM	0,7655	0,6756	0,7103

3. Base Parkinsons

Para a base *Parkinsons*, na tabela 4, observa-se que o modelo proposto teve AUC e acurácia abaixo novamente do modelo SVM e também do modelo MLP. A abordagem proposta ganha novamente dos modelos Gaussianos e do ELM, que foi o que apresentou pior resultado entre os demais. Assim, comprova-se mais uma vez a eficácia do método como classificador além de validar sua eficiência como modelo Gaussiano sem a necessidade de se ajustar parâmetros.

5. CONCLUSÕES

Pode-se concluir que a teoria de conjuntos dominantes aplicada ao grafo de Gabriel foi adequada para encontrar os centros c e os raios σ das Gaussianas das funções RBF (kernel) de maneira autônoma, ou em outras palavras, sem a necessidade de se arbitrar estes parâmetros. Os testes demonstraram que para problemas de classificação de diferentes dimensões o método apresentou resultados próximos dos métodos clássicos ficando atrás apenas do SVM conhecido pela sua robustez. Esta constatação foi feita utilizando uma base sintética e três bases reais. No *benchmarking* mostrado na tabela 1, o método apresentou melhores resultados que outras abordagens Gaussianas que calculam os parâmetros c e σ com arbitragem. Este resultado demonstra a eficiência do método, que também é Gaussiano, validando-o como classificador. No *benchmarking* com os classificadores clássicos, mostrado na tabela 4, o método apresentou resultados próximos das outras abordagens do pacote Caret que usam processos de validação cruzada e ajustes de parâmetros para obter os melhores resultados. De fato, a metodologia perde da SVM que é considerado um ótimo classificador junto à literatura. Por fim, o artigo conseguiu demonstrar que o método pode ser aplicado em problemas de classificação servindo como comparativo para estudos futuros de métodos que utilizam funções de kernel radiais em problemas não linearmente separáveis.

AGRADECIMENTOS

O presente trabalho foi realizado com o apoio financeiro da CAPES, CNPq e FAPEMIG - Brasil.

REFERÊNCIAS

- Berg, M., Cheong, O., Kreveld, M., and Overmars, M. (2008). *Computational Geometry: Algorithms and Applications*. Springer-Verlag TELOS, Santa Clara, CA, USA., third edition.
- Frey, B.J. and Dueck, D. (2007). Clustering by passing messages between data points. *science*, 315(5814), 972–976.
- Haykin, S. (2005). *Neural Networks, A Comprehensive Foundation, Second Edition*. Pearson Education (Singapore) Pte. Ltd., Indian Branch, 482 F.I.E. Patparganj, Delhi 110092, India., second edition.
- Haykin, S. (2009). *Neural networks and learning machines*. Copyright ©2009 by Pearson Education, Inc., Upper Saddle River, New Jersey 07458, third edition.
- Rousseeuw, P.J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20, 53–65.
- Sousa, H.V. (2015). *Um estudo sobre conjuntos cliques-dominantes em grafos*. Dissertação de mestrado, Universidade Estadual de Campinas. Instituto de Computação, Programa de Pós-Graduação em Ciência da Computação, Campinas, São Paulo.
- Souza, P.V.C., Torres, L.C., Braga, A.P., and Coelho, F.G.F. (2014). Análise de algoritmos de maximização de margem baseados em grafos de gabriel e fecho afim.
- Torres, L.C.B. (2012). *Uma nova Abordagem Baseada em Margem Para Seleção de Modelos Neurais*. Dissertação de mestrado, UFMG, Belo Horizonte.
- Torres, L.C.B. (2016). *Classificador por Arestas de Suporte (CLA): Métodos de Aprendizado Baseados em Grafos de Gabriel*. Tese de doutorado, UFMG, Belo Horizonte.
- Wanderley, B.F.M. (2013). *Estudos em Estimção de Densidade por Kernel: Métodos de Seleção de Características e Estimção do Parâmetro Suavizador*. Tese de doutorado, UFMG, Belo Horizonte.
- Zhang, W. and King, I. (2002). *A study of the relationship between support vector machine and gabriel graph*. In *Neural Networks. IJCNN '02. Proceedings of the 2002 International Joint Conference on*, vol. 1.