# MÁQUINAS DE VETOR DE SUPORTE INTERVALAR

Adriana Takahashi*, Adrião D. D. Neto*, Benjamín R. C. Bedregal*

*Universidade Federal do Rio Grande do Norte
Caixa Postal 1524 - Campus Universitário Lagoa Nova, CEP 59072-970
Natal, RN, Brasil

Emails: adriana@dca.ufrn.br, adriao@dca.ufrn.br, benjamin@dimap.ufrn.br

**Abstract**— Support vector machines has been successfully applied to a number of research and applications, however, some cases is not easy to classify correctly and with precision the input patterns in pattern recognition technique and to find the optimal hyperplane, then the interval theory is proposal for cases the incomplete input patterns. This paper proposed a version of SVM using interval theory (SVMI), interval support vector machines. The objective of SVMI is to control the incomplete input patterns and to find support vectors of optimal hyperplane to separation.

**Keywords**— SVM, interval.

**Resumo**— A eficiência das máquinas de vetores-suporte no aprendizado de máquinas tem levado ao desenvolvimento de muitas pesquisas e aplicações associadas, porém, em alguns casos nem sempre é fácil classificar com precisão um determinado padrão entre duas ou mais classes, para problemas de classificação de padrões, e uma vez que, para encontrar o hiperplano de separação ótimo está relacionado diretamente aos dados de entrada aos vetores-suporte, então a teoria intervalar é proposta para casos onde os padrões de entrada não possuam características que modelem alguma classe. Este artigo desenvolve uma nova abordagem para SVM, utilizando SVM associado com a teoria intervalar (SVMI), as máquinas de vetor de suporte intervalares. O objetivo da SVMI é controlar as informações dos padrões de entrada para encontrar vetores-suporte de um hiperplano de separação ótimo quando houver incertezas nos dados ou imprecisões contidas no conjunto de treinamento e controle dos erros computacionais gerados no processo de treinamento da máquina.

**Keywords**— SVM, intervalar.

## 1 Introdução

As máquinas de vetores-suporte (SVM - Support Vector Machines) tem atraído muita atenção nos últimos anos devido a sua eficiência em aplicações que requerem aprendizado de máquina e por estar bem fundamentado na teoria de aprendizado estatístico (Stitson et al., 1996; Pontil and Verri, 1997). As SVMs foram desenvolvidas inicialmente para resolver problemas de classificação, Burges (Burges, 1998) apresenta um tutorial sobre SVM que trata de problemas de classificação de padrões, e Stitson e autores (Stitson et al., 1996; Hearst, 1998) mostram problemas de regressão, fazendo da SVM uma abordagem abrangente para diversas aplicações que envolvem problemas de modelagem de dados empíricos.

As SVMs foram concebidas através de uma formulação que busca a minimização do risco estrutural e utilizam vetores-suporte para achar hiperplanos ótimos, responsáveis para a separação ótima de classes no espaço de entrada.

Mesmo com a eficiência mostrado em pesquisas relacionadas com SVM existem alguns casos onde a SVM não determina otimamente a separação entre classes distintas através do hiperplano ótimo. Nesses casos, o uso da teoria intervalar tem sido bem aceita, pois, desenvolve a máquina nos pontos onde não é possível determinar com precisão a pertinência de padrões do conjunto de treinamento.

Este artigo apresenta uma formalização da SVM para oferecer uma introdução e evidenciar detalhes relevantes para a aplicação da teoria intervalar (Acióly, 1991; Kreinovich et al., 1998; Kulisch, 1982; Kulisch and Miranker, 1981) junto com a SVM.

## 2 Matemática Intervalar

A representação de intervalos no conjunto dos números reais $\mathbb{R}$ é denotado pelo par ordenado de números reais $X = [\underline{x}; \overline{x}]$, tal que $\underline{x} \leq \overline{x}$, e $\underline{x}$ e $\overline{x} \in \mathbb{R}$, e se $\mathbb{R}$ representa o conjunto de todos os números reais, então, $X = \{x \in \mathbb{R} \mid \underline{x} \leq x \leq \overline{x}\}$.

Considere as descrições deste estudo as letras maiúsculas como intervalos do conjunto de intervalos reais, por exemplo, seguindo a definição acima, a letra $X$ representa o intervalo do conjunto intervalar $\mathbb{IR}$, onde, $\underline{x}$ é denominado de ínfimo e $\overline{x}$ denominado de supremo.

A representação de um número real exato é dado como $X = [\underline{x}; \overline{x}]$, em que, $\underline{x} = \overline{x}$, ou seja, se $X = 4$, então $X = [4; 4]$. Este tipo de intervalo é chamado de intervalo degenerado.

### 2.1 Operações Aritméticas Intervalares

Sejam $X = [\underline{x}; \overline{x}]$ e $Y = [\underline{y}; \overline{y}]$, onde, $X$ e $Y \in \mathbb{IR}$. As operações aritméticas, tais como, *adição*, *subtração*, *multiplicação* e *divisão* em $\mathbb{IR}$ são definidas sobre os extremos de seus intervalos, (Moore, 1966).

1. Adição Intervalar:
   $X + Y = [\underline{x} + \underline{y}; \overline{x} + \overline{y}]$

2. Pseudo Inverso Aditivo Intervalar:
   $-X = [-\overline{x}; -\underline{x}]$

3. Subtração Intervalar:
   $X - Y = [\underline{x} - \overline{y}; \overline{x} - \underline{y}]$

4. Multiplicação Intervalar:
   $X * Y = [min\{\underline{x}\underline{y}; \overline{x}\underline{y}; \underline{x}\overline{y}; \overline{xy}\};$
   $max\{\underline{x}\underline{y}; \overline{x}\underline{y}; \underline{x}\overline{y}; \overline{xy}\}]$

5. Pseudo Inverso Multiplicativo Intervalar: $0 \notin X$
   $X^{-1} = \frac{1}{X} = [\frac{1}{\overline{x}}; \frac{1}{\underline{x}}]$

6. Divisão Intervalar: $0 \notin Y$
   $\frac{X}{Y} = [min\{\frac{\underline{x}}{\overline{y}}; \frac{\overline{x}}{\overline{y}}; \frac{\underline{x}}{\underline{y}}; \frac{\overline{x}}{\underline{y}}\}; max\{\frac{\underline{x}}{\overline{y}}; \frac{\overline{x}}{\overline{y}}; \frac{\underline{x}}{\underline{y}}; \frac{\overline{x}}{\underline{y}}\}]$

7. Quadrado Intervalar:
   $$X^2 = \begin{cases} [\underline{x}^2; \overline{x}^2], \text{ se } 0 \leq \underline{x} \\ [\overline{x}^2; \underline{x}^2], \text{ se } \overline{x} \leq 0 \\ [0, \max\{\underline{x}^2; \overline{x}^2\}], \text{ senão} \end{cases}$$

### *2.2 Propriedades Algébricas Intervalares*

Sejam $X$, $Y$, $Z \in \mathbb{IR}$. As propriedades algébricas para as operações anteriores são, *fechamento*, *comutativa*, *associativa*, *elemento neutro*, *subdistributiva*, e *monotônica*.

1. Fechamento:
   - Se $X$, $Y \in \mathbb{IR}$, então $X + Y \in \mathbb{IR}$
   - Se $X$, $Y \in \mathbb{IR}$, então $X * Y \in \mathbb{IR}$
2. Comutativa:
   - $X + Y = Y + X$
   - $X * Y = Y * X$
3. Associativa:
   - $X + (Y + Z) = (X + Y) + Z$
   - $X * (Y * Z) = (X * Y) * Z$
4. Elemento Neutro:
   - $X + [0; 0] = [0; 0] + X = X$
   - $X * [1; 1] = [1; 1] * X = X$
5. Subdistributiva:
   $X * (Y + Z) \subseteq (X * Y) + (X * Z)$
6. Inclusão Monotônica:
   Sejam $X$, $Y$, $Z$ e $W \in \mathbb{IR}$, tais que, $X \subseteq Z$ e $Y \subseteq W$.
   - $X + Y \subseteq Z + W$
   - $-X \subseteq -Z$
   - $X - Y \subseteq Z - W$
   - $X * Y \subseteq Z * W$.
   - $\frac{1}{X} \subseteq \frac{1}{Z}$, se $0 \notin Z$
   - $\frac{X}{Y} \subseteq \frac{Z}{W}$, se $0 \notin W$

### *2.3 Ordem Intervalar*

Na literatura encontramos diversas formas de definição de ordens (parciais) para intervalos. As mais conhecidas são, ordem de Moore (Moore, 1966), ordem de Kulisch & Miranker (Kulisch and Miranker, 1981), ordem da Informação (Acióly, 1991) e ordem da Teoria dos Conjuntos.

1. Ordem de Moore:
   $X < Y = [\underline{x}; \overline{x}] < [\underline{y}; \overline{y}] \Leftrightarrow \overline{x} < \underline{y}$

2. Ordem de Kulisch-Miranker:
   $X \leq Y = [\underline{x}; \overline{x}] \leq [\underline{y}; \overline{y}] \Leftrightarrow \underline{x} \leq \underline{y}$ e $\overline{x} \leq \overline{y}$

3. Ordem da Teoria dos Conjuntos:
   $X < Y = [\underline{x}; \overline{x}] \leq [\underline{y}; \overline{y}] \Leftrightarrow [\underline{x}; \overline{x}] \subseteq [\underline{y}; \overline{y}] \Leftrightarrow \underline{y} \leq \underline{x}$ e $\overline{x} \leq \overline{y}$. Logo $[x_1; x_2] \leq [y_1; y_2]$.

4. Ordem da Informação:
   $X \leq Y = [\underline{x}; \overline{x}] \leq [\underline{y}; \overline{y}] \Leftrightarrow [\underline{y}; \overline{y}] \subseteq [\underline{x}; \overline{x}] \Leftrightarrow \underline{x} \leq \underline{y}$ e $\overline{y} \leq \overline{x}$. Logo, $[\underline{x}; \overline{x}] \leq [\underline{y}; \overline{y}]$.

### *2.4 Função Intervalar*

Sejam $F : \mathbb{IR} \rightarrow \mathbb{IR}$ e $f : \mathbb{R} \rightarrow \mathbb{R}$, então $F$ representa $f$ e $f$ é representado por $F$, logo $f \subseteq F$, se:

$$\forall X \in \mathbb{IR}, \forall x \in X \text{ temos que } f(x) \in F(X)$$

A extensão dessa definição de função para $X \subset \mathbb{IR}^m$ e $Y \subset \mathbb{IR}^n$ é feita de forma natural.

### *2.5 Derivada Intervalar*

Seja $F : \mathbb{IR} \rightarrow \mathbb{IR}$ contínua cujas funções reais mínimas e máximas são deriváveis, é dito que a derivada de $F$ em relação a variável $X \in \mathbb{IR}$ é a função intervalar $F'[a, b]$, tal que, (Lyra, 2003):

$$\begin{aligned} \frac{dF}{dX} = \; & [min\{\{\frac{df_{Rmin}}{dx}/x \in X\} \bigcup \{\frac{df_{Rmax}}{dx}/x \in X\}\}, \\ & max\{\{\frac{df_{Rmin}}{dx}/x \in X\} \bigcup \{\frac{df_{Rmax}}{dx}/x \in X\}\}] \end{aligned} \quad (1)$$

### *2.6 Matriz Intervalar*

Uma matriz intervalar $A = a_{ij}$ de ordem $m$ por $n$ é dita intervalar se cada elemento $a_{ij}$ for um intervalo (Lyra, 2003).

Seja $A = a_{ij} \in \mathbb{R}$ de ordem $m \times n$, então, a matriz $A$, formada pelos intervalos $a_{ij}$ é uma matriz intervalar de índice $i$ representando a linha e $j$ a coluna da matriz intervalar.

### *2.7 Vetor Intervalar*

Seja $A = a_{ij}$ uma matriz intervalar de ordem $m \times n$, se $m = 1$ ou $n = 1$, então a matriz $A$ é denominada de vetor intervalar.

## 3 Máquinas de Vetores-Suporte Intervalares Linearmente Separáveis

Máquinas de vetores-suporte é uma técnica de aprendizado de máquina baseada na teoria do aprendizado estatístico (Haykin, 2001) para aprendizado supervisionado.

A tarefa de classificar padrões é feito através da função:

$$f(\mathbf{x}) = sgn(\sum_{i=1}^{N_{VS}} d_i \alpha_i \mathbf{x_i}^T \mathbf{x_i} + b) \qquad (2)$$

onde, $x_i \in \mathbb{R}^n$ é o vetor de entrada de n-dimensão, $d_i \in \{-1, 1\}$ é a classe que pertence $x_i$, e $\alpha_i$ e $b$ são parâmetros da função objetivo encontrados durante o treinamento através da resolução de um problema de otimização.

Para uma SVMI, considere uma amostra de treinamento $\{(\mathbf{X}_i, D_i)\}_{i=1}^N$, onde, $\mathbf{X}_i$ é intervalo que representa o padrão de entrada para o $i$-ésimo exemplo e $D_i = [\underline{d}_i; \overline{d}_i] \in \{[+1;+1], [-1;-1]\}$ é a resposta desejada do $i$-ésimo padrão de entrada.

A equação de uma superfície de decisão na forma de hiperplano que realiza a separação entre as classes é definida como:

$$\mathbf{W}^T \mathbf{X} + B = [0; 0] \qquad (3)$$

onde, $\mathbf{X}$ é um vetor intervalar de entrada, $\mathbf{W}$ é um vetor intervalar peso ajustável, $B$ é o bias e $[0; 0]$ um intervalo degenerado.

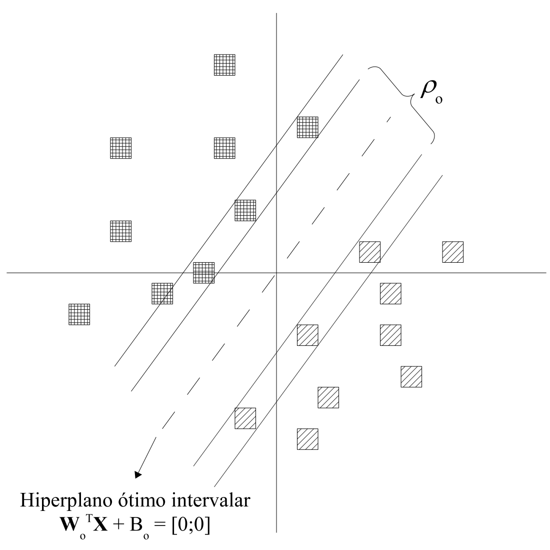


Figura 1: Hiperplano ótimo para padrões intervalares linearmente separáveis

O conjunto de treinamento $\{(\mathbf{X}_i, D_i)\}_{i=1}^N$ é dito linearmente separável se existir um vetor intervalar $W$ e um intervalo $B$ que satisfaça:

$$\begin{cases} \mathbf{W}_o^T \mathbf{X}_i + B_o \geq [+1;+1], D_i = [+1;+1] \\ \mathbf{W}_o^T \mathbf{X}_i + B_o \leq [-1;-1], D_i = [-1;-1] \end{cases} \qquad (4)$$

que é equivalente a:

$$D_i(\mathbf{W}^T \mathbf{X}_i + B) \geq [1; 1] \qquad (5)$$

onde, o par $(\mathbf{W}, B)$ define o hiperplano de separação da equação (3).

Para obter o hiperplano ótimo, ou seja, o hiperplano de máxima margem entre as classes é necessário encontrar a distância de um intervalo $\mathbf{X}_i$ com o hiperplano de separação $(\mathbf{W}, B)$. Considerando a função discriminante:

$$G(\mathbf{X}) = \mathbf{W}_o^T \mathbf{X}_i + B_o \qquad (6)$$

onde, $G(\mathbf{X})$ fornece uma medida algébrica da distância de $\mathbf{X}$ até o hiperplano, podendo ser expresso também da seguinde forma:

$$\mathbf{X} = \mathbf{X}_p + R \frac{\mathbf{W}_o}{\parallel \mathbf{W}_o \parallel} \qquad (7)$$

onde, $\mathbf{X}_p$ é a projeção normal de $\mathbf{X}$ sobre o hiperplano ótimo, e $R$ é a distância.

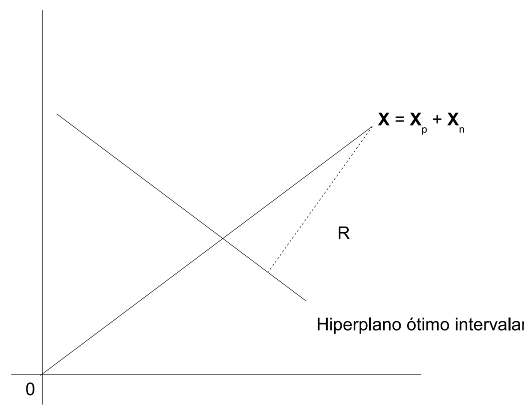


Figura 2: Interpretação da distância de $\mathbf{X}$ até p hiperplano ótimo.

O vetor normal é dado por:

$$\mathbf{X}_n = R \frac{\mathbf{W}_o}{\parallel \mathbf{W}_o \parallel} \qquad (8)$$

Para o caso de separação ótima, por definição, $G(\mathbf{X}_p) = [0; 0]$, resultando:

$$\begin{aligned} G(\mathbf{X}) = &\ \mathbf{W}_o^T \mathbf{X} + B_o \\ = &\ \mathbf{W}_o^T (\mathbf{X}_p + \mathbf{X}_n) + B_o \\ \subseteq &\ \mathbf{W}_o^T \mathbf{X}_p + B_o + \mathbf{W}_o^T \mathbf{X}_n \\ = &\ [0; 0] + \mathbf{W}_o^T \mathbf{X}_n \\ = &\ \mathbf{W}_o^T \frac{\mathbf{W}_o}{\|\mathbf{W}_o\|} R \\ = &\ R \frac{\|\mathbf{W}_o\|^2}{\|\mathbf{W}_o\|} \\ = &\ R \parallel \mathbf{W}_o \parallel \end{aligned} \qquad (9)$$

Logo,

$$R = \frac{G(\mathbf{X})}{\parallel \mathbf{W}_o \parallel} \qquad (10)$$

Para encontrar os parâmetros $\mathbf{W}_o$ e $B_o$ para o hiperplano ótimo dado um conjunto de treinamento, as restrições da equação (15) devem ser satisfeitas. Os pontos intervalares $(\mathbf{X}_i, D_i)$ satisfeito no sinal de igualdade da equação (15) são chamados de **vetores-suporte** intervalares.

Considerando um vetor de suporte intervalar $\mathbf{X}^(s)$ temos:

$$G(\mathbf{X}^{(s)}) = \mathbf{W}_o^T\mathbf{X}^{(s)} + B_o \pm [1;1], D^{(s)} = \pm[1;1] \quad (11)$$

Da equação (10) a distância do vetor de suporte intervalar até o hiperplano ótimo:

$$R = \begin{cases} +\frac{[1;1]}{\|\mathbf{W}_o\|} \text{ se } D^{(s)} = +[1;1] \\ -\frac{[1;1]}{\|\mathbf{W}_o\|} \text{ se } D^{(s)} = -[1;1] \end{cases} \quad (12)$$

Considerando que $\rho$ represente o valor intervalar ótimo da margem de separação entre as duas classes do conjunto de treinamento, então:

$$\begin{aligned} \rho = & \; 2R \\ = & \; \frac{2}{\|\mathbf{W}_o\|} \end{aligned} \quad (13)$$

Da equação (13) temos que, maximizar a margem de separação entre as classes é equivalente a minimizar a norma do vetor intervalar $\mathbf{W}$ com respeito as restrições.

O hiperplano da equação (3) é único e de máxima separação entre as classes.

Para encontrar o hiperplano ótimo utilizando o conjunto de treinamento $\{(\mathbf{X}_i, D_i)\}_{i=1}^N$ e que satisfaça as restrições é necessário encontrar os parâmentros $\mathbf{W}$ e $B$.

O hiperplano de separação ótimo minimiza a função custo intervalar:

$$\Phi = \frac{1}{2}\mathbf{W}^T\mathbf{W} \quad (14)$$

sujeito as restrições:

$$D_i(\mathbf{W}^T\mathbf{X}_i + B) \geq [1;1]$$

Este é um problema de otimização que pode ser resolvido através do método de multiplicadores de Lagrange adaptado para intervalar:

$$\begin{aligned} J(\mathbf{W}, B, \alpha) = & \; [\tfrac{1}{2};\tfrac{1}{2}]\mathbf{W}^T\mathbf{W} - \\ & \textstyle\sum_{i=1}^N \alpha_i(D_i(\mathbf{W}_i^T\mathbf{X}_i + B) - [1;1]) \\ \subseteq & \; [\tfrac{1}{2};\tfrac{1}{2}]\mathbf{W}^T\mathbf{W} - \\ & \textstyle\sum_{i=1}^N \alpha_i D_i \mathbf{W}_i^T\mathbf{X}_i + \\ & \textstyle\sum_{i=1}^N \alpha_i D_i B - \sum_{i=1}^N \alpha_i \end{aligned} \quad (15)$$

onde, $\alpha_i$ são intervalos representando os multiplicadores de Lagrange.

A função lagrangiana tem que ser minimizada com respeito a $\mathbf{W}$, $B$ e maximizada com respeito a $\alpha_i \geq 0$. Assim, diferenciando $L(\mathbf{W}, B, \alpha)$ em relaçõ a $\mathbf{W}$ $B$ temos as seguintes condições:

$$\begin{aligned} \frac{\partial L(\mathbf{W},B,\alpha)}{\partial B} = [0;0] \Rightarrow \textstyle\sum_{i=1}^N \alpha_i D_i = [0;0] \\ \frac{\partial L(\mathbf{W},B,\alpha)}{\partial \mathbf{W}} = [0;0] \Rightarrow \mathbf{W} = \textstyle\sum_{i=1}^N \alpha_i D_i \mathbf{X}_i \end{aligned} \quad (16)$$

Substituindo as condições (16) em (15) temos:

$$\begin{aligned} J(\mathbf{W}, B, \alpha) = & \; [\tfrac{1}{2};\tfrac{1}{2}]\mathbf{W}^T\mathbf{W} - \\ & \textstyle\sum_{i=1}^N \alpha_i D_i \mathbf{W}_i^T\mathbf{X}_i + \\ & \textstyle\sum_{i=1}^N \alpha_i D_i B - \sum_{i=1}^N \alpha_i \\ = & \; [\tfrac{1}{2};\tfrac{1}{2}]\textstyle\sum_{i,j=1}^N \alpha_i D_i \mathbf{X}_i \alpha_j D_j \mathbf{X}_j - \\ & \textstyle\sum_{i,j=1}^N \alpha_i D_i \mathbf{X}_i \alpha_j D_j \mathbf{X}_j - [0;0] + \\ & \textstyle\sum_{i=1}^N \alpha_i \\ = & \; -[\tfrac{1}{2};\tfrac{1}{2}]\textstyle\sum_{i,j=1}^N \alpha_i D_i \mathbf{X}_i \alpha_j D_j \mathbf{X}_j \\ & + \textstyle\sum_{i=1}^n \alpha_i \end{aligned} \quad (17)$$

Fazendo a função objetivo $J(\mathbf{W}, B, \alpha) = Q(\alpha)$ temos:

$$Q(\alpha) = \sum_{i=1}^N \alpha_i - [\frac{1}{2};\frac{1}{2}]\sum_{i,j=1}^N \alpha_i\alpha_j D_i D_j \mathbf{X}_i^T\mathbf{X}_j \quad (18)$$

Logo, o problema dual é dado por:

$$\begin{aligned} & \text{Maximizar:} \\ & Q(\alpha) = \textstyle\sum_{i=1}^N \alpha_i - [\tfrac{1}{2};\tfrac{1}{2}]\sum_{i,j=1}^N \alpha_i\alpha_j D_i D_j \mathbf{X}_i^T\mathbf{X}_j \\ & \text{Sujeito as restrições:} \\ & \begin{cases} (1) \; \alpha_i \geq [0;0], i = 1, ..., N \\ (2) \; \textstyle\sum_{i=1}^N \alpha_i D_i = [0;0] \end{cases} \end{aligned} \quad (19)$$

Ao encontrar os multiplicadores de Lagrange é possível calcular os pesos intervalares ótimos:

$$\mathbf{W}_o = \sum_{i=1}^N \alpha_{oi} D_i \mathbf{X}_i \quad (20)$$

O valor do bias ótimo $B_o$ é encontrado utilizando os pesos ótimos $\mathbf{W}_o$ encontrados na equação (20) e descrito como:

$$B_o = [1;1] - \mathbf{W}_o^T\mathbf{X}^{(s)} \text{ para } D^{(s)} = [1;1] \quad (21)$$

## 4 Conclusões

Neste artigo foi apresentado conceitos e teorias para o desenvolvimento do formalismo de uma nova abordagem para SVM, SVM intervalar, para o caso linearmente separável. Foi utilizado em todo o formalismo da SVM intervalar conceitos da matemática intervalar, desde os dados iniciais, conjunto de treinamento representados por intervalos, até todo o processamento de operações estendidas dos conjuntos dos reais para intervalares.

Conclue-se que, esta nova abordagem, SVM intervalar, pode ser tratada em problemas que não exijam resultados rígidos, ou seja, o hiperplano de separação ótimo é encontrado para a classificação das classes, porém, tratando com a imprecisão dos dados iniciais, um conjunto de padrões imprecisos são encontrados no resultado da classificação para uma posterior análise dependente do problema analisado


## Agradecimentos

Este trabalho contou com o apoio do CNPq.

## Referências


Acióly, B. M. (1991). *Fundamentação Computacional da Matemática Intervalar*, PhD thesis, Universidade Federal do Rio Grande do Sul.

Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition, *Data Mining and Knowledge Discovery* **2**: 121–167.

Haykin, S. (2001). *Redes Neurais: Princípios e prática*, Bookman.

Hearst, M. A. (1998). Support vector machines, *IEEE Inteligent Systems* **13**(4): 18–28.

Kreinovich, V., Rohn, A. L. J. and Kahl, P. (1998). *Computational Complexity and Feability of Data Processing and Interval Computations*, Kluwer Academic Publishers, Canadá.

Kulisch, U. W. (1982). Computer arithmetic and programing languages, *ACM* **13**: 176–182.

Kulisch, U. W. and Miranker, W. L. (1981). *Computer Arithmetic Theory and Pratice*, Academin Press.

Lyra, A. (2003). *Uma Fundamentação Matemática para o Processamento de Imagens Digitais Intervalares*, Thesis(Ph.D.), Universidade Federal do Rio Grande do Norte, Natal.

Moore, R. E. (1966). *Interval Analysis*, Pretice Halls, New Jersey.

Pontil, M. and Verri, A. (1997). Proprieties of support vector machines, *Technical report*, Massachusetts Institute of Technology.

Stitson, M. O., Weston, J. A. E., Gammerman, A., Vovk, V. and Vapnik, V. (1996). Theory oh support vector machines, *Technical report*, University of London.