

# ESTUDO COMPARATIVO DE ANÁLISE E CLASSIFICAÇÃO DE BASE DE DADOS UTILIZANDO REDES NEURAIS MLP E MÁQUINAS DE VETORES DE SUPORTE SVM

RICARDO BRITO<sup>1</sup>

1. SENAI CIMATEC

Av. Orlando Gomes, 1845 - Piatã, Salvador - BA, CEP 41650-010

E-mail: [ricardobrito\\_348@hotmail.com](mailto:ricardobrito_348@hotmail.com)

**Abstract** - This work aims to show the results of the use of two neural network implementation techniques to analyze and classify a database (named Car Evaluation) for automotive vehicles. In particular, this database has information about several characteristics that can be used as evaluation criteria (price, comfort, technology) for the quality of a vehicle. In the first part of this work, we will present the database and its peculiarities, as well as the concepts related to support vector machine (SVM) and multi-layer perceptron (MLP) techniques. In the second part, the implementation of these techniques is performed, verifying the performance and analyzing the results of each of them regarding the dynamics and precision in the data classification. In the last part, the final considerations of the techniques are presented in front of the obtained results.

**Keywords** - Car Evaluation, UCI Machine, MLP, SVM, Neural Networks.

**Resumo** - Este trabalho objetiva mostrar os resultados da utilização de duas técnicas de implementação de redes neurais para análise e classificação de um banco de dados (denominado Car Evaluation) referente a veículos automotivos. Em particular, este banco de dados possui informações acerca de diversas características que podem ser utilizadas como critérios para avaliação (preço, conforto, tecnologia) quanto à qualidade de um veículo. Na primeira parte deste trabalho, será apresentado o banco de dados e suas particularidades, bem como os conceitos referentes às técnicas de máquina de vetor de suporte (SVM) e perceptron multi-camadas (MLP). Na segunda parte, é realizada a implementação destas técnicas, verificando-se o desempenho e analisando os resultados de cada uma destas quanto à dinâmica e precisão na classificação dos dados. Na última parte, são apresentadas as considerações finais das técnicas frente aos resultados obtidos.

**Palavras-chave** - Car Evaluation, UCI Machine, MLP, SVM, Redes Neurais.

## 1. INTRODUÇÃO

Os estudos das Redes Neurais Artificiais (RNA) se iniciaram por volta da década de 40, quando o neurofisiologista Warren McCulloch e o matemático Walter Pitts fizeram uma analogia entre células nervosas e o processo eletrônico, divulgado por eles através de um artigo cujo título foi: A Logical Calculus of the Ideas Immanent in Nervous Activity.

Durante as décadas seguintes, diversos trabalhos e estudos foram realizados na área da inteligência artificial, a qual teve sua primeira conferência internacional em 1956, onde foi apresentado o modelo de uma rede neural artificial pelo pesquisador da IBM Nathaniel Rochester. Em 1959 foi batizada, por Frank Rosenblatt, a rede de perceptron, uma rede com múltiplos neurônios do tipo discriminadores lineares. Muitas descobertas e diversos modelos foram criados ao longo de anos até chegar em 1987, quando se teve a Primeira Conferência de Redes Neurais e também a formação da Sociedade Internacional de Redes Neurais (International Neural Networks Society - INNS). A partir deste momento, as instituições começaram a financiar diversos estudos e programas educacionais na área.

As tecnologias das Redes Neurais Artificiais surgiram devido à demanda em solucionar problemas relacionados ao reconhecimento de padrões baseadas em informações previamente conhecidas. As redes neurais são baseadas no comportamento dos neurônios humanos, as quais tendem a aprender e se adaptar às diversas situações ou problemas que são impostos à ela. Uma Rede Neural Artificial pode ser vista como um conjunto de unidades, conhecidas como neurônios artificiais, interconectadas entre si, com cada uma dessas unidades, possuindo uma pequena quantidade de memória. Dessa forma, busca-se assemelhar ao máximo as RNA ao comportamento do cérebro humano, com o intuito de fazer com que elas possam desempenhar funções similares futuramente.

Neste trabalho serão abordados dois modelos de aprendizados aplicados a um banco de dados com informações previamente estabelecidas. Os treinamentos para essas funções serão do tipo supervisionados, onde tem-se a entrada dos dados e a saída que se espera encontrar para cada situação. Os modelos a serem trabalhados são conhecidos como Multi-Layer Perceptron (MLP) e o Support Vector Machine (SVM). O trabalho se resume a um teste de classificação em que, a partir de um banco de dados e dos algoritmos acima treinados e validados, serão testados valores para

entender, estudar e comparar seus comportamentos. O banco de dados escolhido é o Car Evaluation (UCI – Machine Learning Repository).

Este trabalho, em seu segundo capítulo, abordará um estudo do banco de dados adotado, juntamente com a apresentação dos conceitos necessários à implementação dos modelos escolhidos.

Serão apresentados, no item 3 deste trabalho, os resultados obtidos através dos testes elaborados para ambos os modelos. Serão feitos comentários e comparações à cerca desses encontrados. Neste cenário, objetiva-se principalmente verificar a eficácia e resposta dos modelos de redes neurais para o tipo de análise proposta neste trabalho, analisando-se conjuntamente as matrizes de confusão (confusion matrix) de cada método.

No item 4 são apresentadas as conclusões referentes aos resultados gerais obtidos, com a aplicabilidade e restrições para os métodos utilizados e proposições de trabalhos futuros.

No item 5, são expressos os agradecimentos para a elaboração e conclusão deste trabalho.

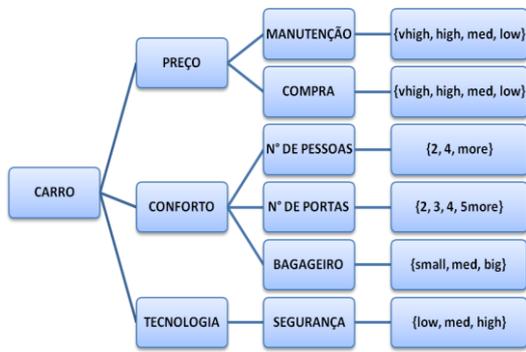
## 2. CARACTERÍSTICAS DO BANCO DE DADOS E DOS MODELOS DE APRENDIZADO

### 2.1 Banco de Dados “Car Evaluation”

O “Car Evaluation Data Set” é uma base de dados extraída do UCI Machine Learning e disponível em [5]. Este conjunto de dados foi inicialmente desenvolvido para modelos matemáticos simples de decisão hierárquica, podendo, segundo [4], ser utilizada também para métodos de testes e descoberta por indução.

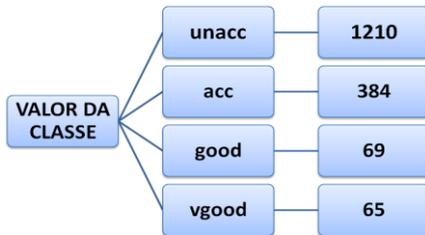
Esta base contém dados de diversos modelos e tipos de carros, sendo multivariável e possuindo seis atributos categorizados (manutenção, compra, número de pessoas, número de portas, bagageiro, segurança), derivados dos conceitos intermediários de preço, conforto e tecnologia. As relações lógicas entre os atributos e conceitos encontram-se ilustrados na figura 1.

Figura 1. Representação esquemática da base de dados car evaluation.



Como resultados dos métodos de classificação apresentados na referência [4], esta base de dados possui quatro classes de resultados, cujas quantidades por classe são apresentadas na figura 2, num total de 1728 instâncias. Estas totalizações parciais, por classe, serão utilizadas como referência ao longo deste trabalho. Outras aplicações desta base de dados em classificação de dados podem ser vistas em [3].

Figura 2. Tipos e valores totalizados para cada classe – base de dados Car Evaluation



## 2.2 Máquinas de Vetores de Suporte - SVM

A Support Vector Machine (SVM) constitui uma técnica de aprendizado do tipo feedforward, ou seja, alimentada para frente. É um algoritmo de aprendizado estatístico, o qual consiste basicamente em, a partir de um conjunto de dados para treinamento, construir um hiperplano como superfície de decisão buscando maximizar a separação entre exemplos positivos e negativos. Esse modelo de aprendizado faz uso de um algoritmo baseado em “vetores de suporte” (os quais são pequenos conjuntos formados a partir dos dados de entrada treinados) e o vetor com os dados de entrada, ou seja, esse modelo consegue prever corretamente as classes de novos dados do mesmo domínio que o aprendizado ocorreu. Esse algoritmo tem como base o método ótimo de Kernel para o qual, em contrapartida, exige grande custo computacional. A SVM pode ser usada para resolver tanto, problemas de regressão não-linear, como problemas de classificação de padrões, sendo neste

último onde vem se obtendo os melhores resultados.

As SVMs possuem sua teoria fundamentada na Teoria de Aprendizado Estatístico (TAE) a qual consiste em estabelecer condições matemáticas para a escolha de um classificador, a partir de um conjunto de treinamento, que venha a apresentar um bom desempenho, evitando situações de sobre-ajustamento e sub-ajustamento [1].

A SVM visa a minimização do erro estrutural, ou seja, busca minimizar um limite superior sobre o erro esperado. Essa é uma das definições proposta pela TAE. A expressão que descreve a minimização do erro estrutural é dada pela relação entre o erro esperado pelo classificador ( $R(f)$ ), o erro empírico ( $R_{emp}(f)$ ) e um termo de capacidade, conforme mostrada na equação (1):

$$R(f) \leq R_{emp}(f) + \sqrt{\frac{h * (\ln(2 * n / h) + 1) - \ln(\theta/4)}{n}} \quad (1)$$

Na Eq.(1), a variável  $h$  representa a dimensão Vapnik-Chervonenkis (VC), a qual é calculada como o tamanho do maior conjunto de pontos que pode ser particionado arbitrariamente, em duas classes, pelas funções de classificação [1]. Quanto maior o valor de  $h$ , mais complexas são estas funções de classificação. O limite, representado pela Eq.(1), é garantido com probabilidade  $1 - \theta$ , em que  $\theta \in [0,1]$ . A variável  $n$  representa a quantidade de exemplos no conjunto de treinamento. O erro esperado pelo classificador ( $R(f)$ ), o qual mede a sua capacidade de generalização é dependente de uma função custo  $c(f(x), y)$  (que relaciona a previsão  $f(x)$  com a saída  $y$  desejada) e de uma distribuição de probabilidade  $P(x, y)$  (a qual descreve a relação entre os dados e seus rótulos), como é mostrado abaixo na equação (2) [2]:

$$R(f) = \int c(f(x), y) dP(x, y) \quad (2)$$

Devido ao desconhecimento da distribuição de probabilidade ( $P(x, y)$ ), este erro não pode ser minimizado. O erro empírico ( $R_{emp}(f)$ ), que mede o desempenho do classificador nos dados de treinamento, está relacionado a uma função custo e é definida pela equação (3):

$$R_{emp}(f) = \frac{1}{n} \sum_{i=1}^n c(f(x_i), y_i) \quad (3)$$

Para este trabalho, nos limitaremos somente a essa forma de escolha de classificador proposta pela Teoria de Aprendizado Estatístico (TAE) que visa a redução do erro estrutural por meio da minimização de um limite superior do erro esperado para o classificador.

Dois casos de grande importância em que as SVMs são aplicadas são os casos Lineares e não Lineares. As SVMs Lineares têm por objetivo a separação de duas classes por meio de uma função que é induzida a partir dos exemplos disponíveis para aprendizagem. Existem diversos classificadores lineares capazes de separar os dados, mas somente um que maximiza a margem de separação, o qual é denominado de hiperplano ótimo. Para as SVMs não Lineares, o conceito aplicado é que não se é possível fazer a separação dos dados, de forma satisfatória, por um hiperplano ótimo. A técnica usada, em problemas não lineares, é o mapeamento do conjunto de treinamento de seu espaço original, referenciado como de entradas, para um novo espaço de maior dimensão, denominado espaço de características (nesse espaço há grande probabilidade dos dados serem linearmente separáveis). São necessárias duas condições para se criar o espaço de características e se obter bons resultados, as quais são: que a transformação seja não linear e a dimensão do espaço de características seja suficientemente alta [2].

Como o mapeamento,  $\gamma$ , do conjunto de dados,  $X$ , para um novo espaço,  $\beta$  ( $\gamma: X \rightarrow \beta$ ), pode ser muito custoso computacionalmente, torna-se necessário o uso de funções específicas denominadas Kernels ( $K$ ) para este procedimento. Um Kernel  $K$  é uma função que recebe dois pontos no espaço de entrada e computa o produto escalar desses dados no espaço no espaço de características [2]. Abaixo seguem a equação (4) de Kernel e a tabela 1, com os Kernels mais utilizados.

$$K(x_i, x_j) = \gamma(x_i) \cdot \gamma(x_j) \quad (4)$$

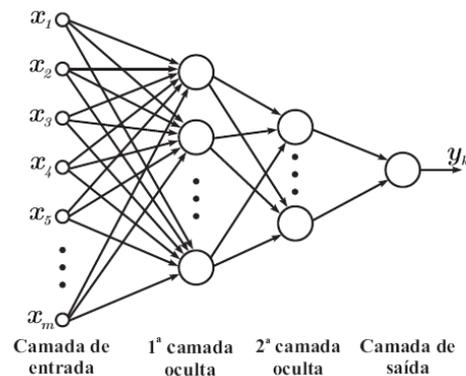
Tabela 1. Funções Kernel mais utilizadas.

Tipo de Kernel	Função $K(x_i, x_j)$	Parâmetros
Polinomial	$(\delta(x_i, x_j) + k)^d$	$\delta, k, d$
Gaussiano	$\exp(-\sigma \ x_i - x_j\ ^2)$	$\sigma$
Sigmoidal	$\tanh(\delta(x_i, x_j) + k)$	$\delta, k$

## 2.3 Redes Neurais do Tipo MLP

O segundo método de classificação escolhido foi a Rede Neural do tipo "Multi-Layer Perceptron", ou Perceptron Multi-Camada, também conhecida como "MLP". Essa arquitetura se diferencia da Perceptron simples por contar com uma ou mais camadas ocultas de neurônios entre os nós de entrada e saída, conforme ilustra a figura 3.

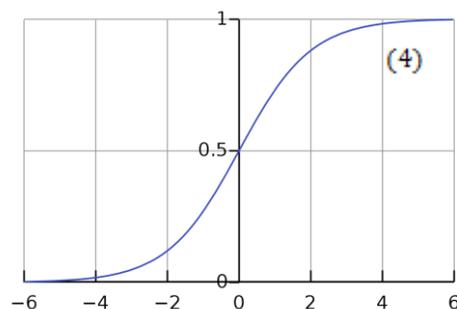
Figura 3. Topologia de exemplo de Rede Neural MLP



Este modelo de rede neural tem sido aplicado, sobretudo, nos problemas de classificação onde se faz necessário reconhecer padrões não-lineares, o que não é possível de forma satisfatória com a rede Perceptron convencional [7].

A função de ativação utilizada foi a função sigmoide, ilustrada na figura 4, enquanto o algoritmo de treinamento escolhido foi o back-propagation. As rotinas de treinamento foram realizadas 10 vezes para cada configuração de rede, com amostras diferentes no vetor de treinamento, a fim de minimizar a aleatoriedade da escolha inicial dos pesos sinápticos da rede neural.

Figura 4 - Função Sigmóide simples



Além disso, foram utilizados para treinamento, teste e validação do aprendizado da rede, respectivamente 40%, 35% e 25% das amostras do banco de dados utilizado.

### 3. IMPLEMENTAÇÃO DOS MODELOS E RESULTADOS OBTIDOS

Nesta etapa do trabalho, serão apresentados os resultados obtidos para cada modelo de rede neural descrito.

#### 3.1. Método SVM

Com base no que foi apresentado no capítulo anterior, implementou-se um código no programa MatLab, fazendo uso da função fitcecoc(), para aplicar o algoritmo da SVM ao banco de dados do Car Evaluation.

Foram aplicados os algoritmos considerando o problema como linear (fez-se uso da função de polinomial de Kernel para o expoente igual a um) e considerando o problema como não linear (aplicou-se a função polinomial de Kernel quadrática – expoente igual a dois).

Para cada uma das situações foram montadas suas respectivas matrizes de confusão e traçadas suas curvas ROCs com ajuda Classification Learner, uma ferramenta presente no MatLab. O estudo abaixo considerou o número de partições na validação cruzada igual a dez. A seguir são mostrados os dados obtidos para a simulação do banco de dados para a SVM Linear, nas figuras 5, 6, 7, 8 e 9.

Figura 5 – Matriz de confusão para SVM Linear

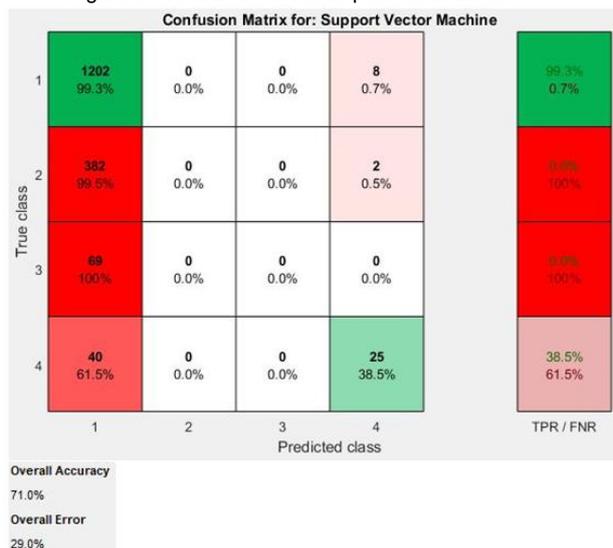


Figura 6: Curva ROC considerando a classe 1 positiva (Área =0,8124)

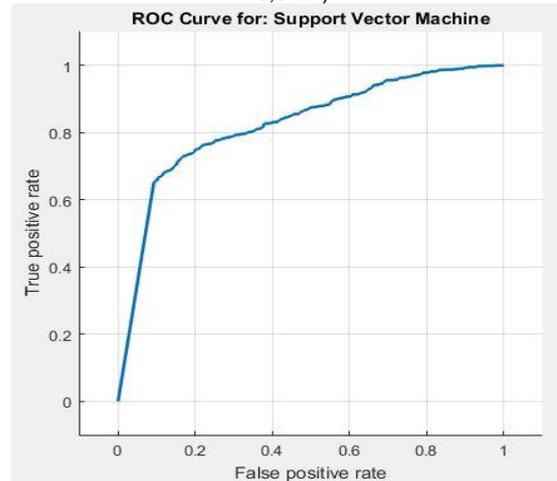


Figura 7: Curva ROC considerando a classe 2 positiva (Área = 0,7728)

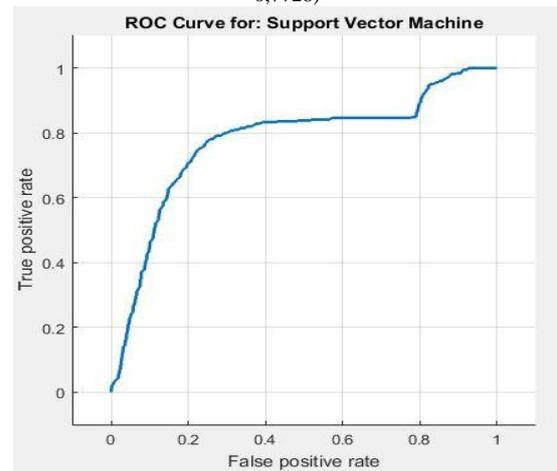


Figura 8: Curva ROC considerando a classe 3 positiva (Área = 0,9688)

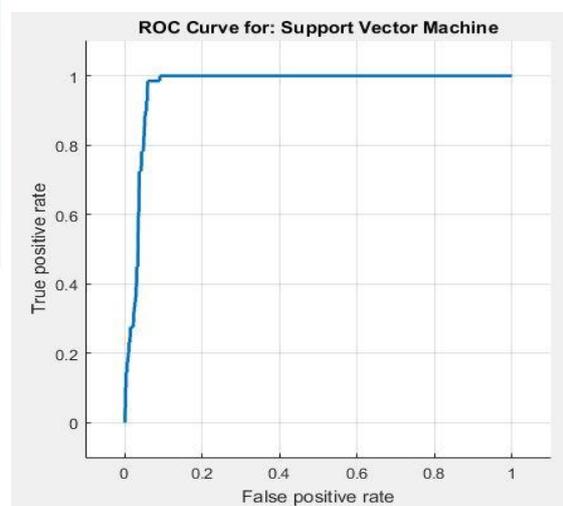
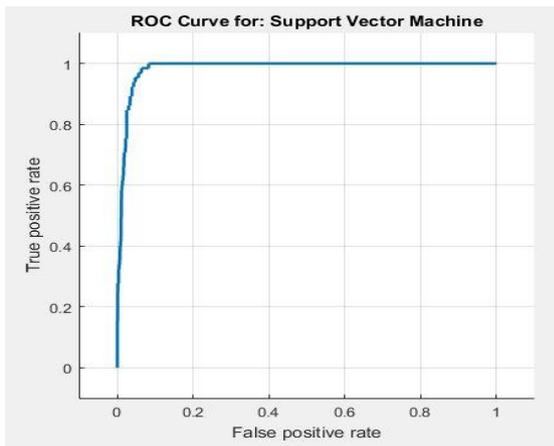


Figura 9: Curva ROC considerando a classe 4 positiva (Área = 0,9852)



Com relação aos dados mostrados acima nas figuras 5, 6, 7, 8, e 9, percebe-se, a partir da matriz de confusão, um erro de 29%, o que torna esse resultado bastante insatisfatório. Pode-se perceber que as classes 2 e 3 não tiveram acertos, ou seja, o classificador mostrou-se muito ineficaz para essas duas situações. As curvas ROCs retratam bem essa situação, como são mostradas nas figuras 6 à 9, onde é possível se perceber a coerência entre elas e as colunas da matriz de confusão. Cada coluna está relacionada a uma curva ROC e a eficiência de cada uma delas está diretamente relacionada com sua área.

Após ser encontrado um resultado insuficiente na aplicação da SVM Linear ao banco de dados utilizado, decidiu-se por se fazer um comparativo com a SVM Quadrática. Os resultados obtidos para sua matriz de confusão e suas curvas ROCs serão ilustrados nas figuras 10, 11, 12, 13 e 14.

Figura 10 – Matriz de confusão para SVM Quadrática

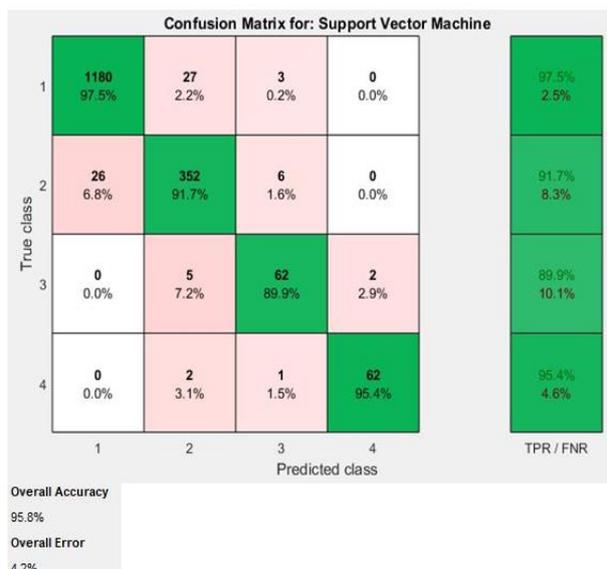


Figura 11: Curva ROC considerando a classe 1 positiva (Área = 0,9948)

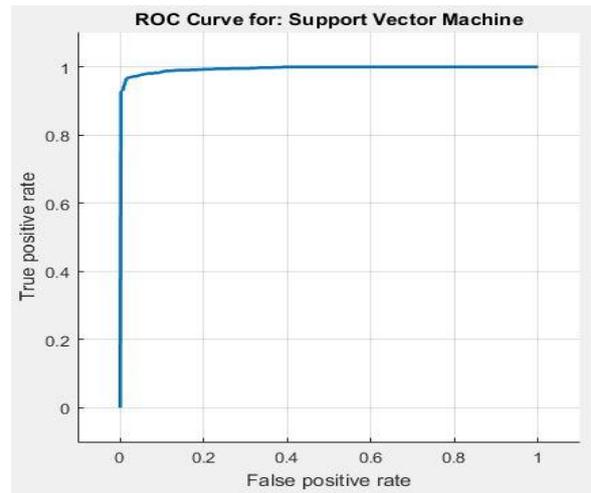


Figura 12: Curva ROC considerando a classe 2 positiva (Área = 0,9905)

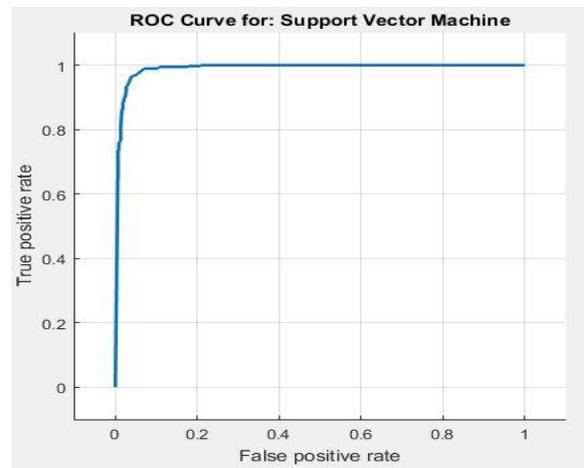


Figura 13: Curva ROC considerando a classe 3 positiva (Área = 0,9974)

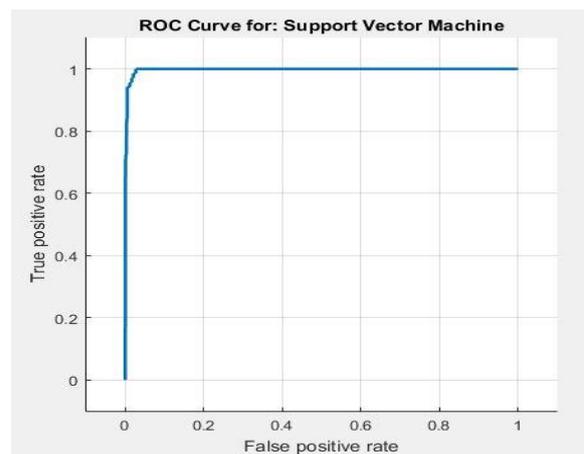
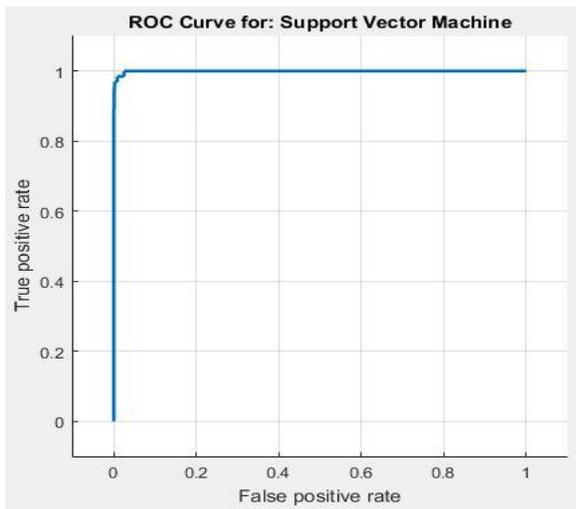


Figura 14: Curva ROC considerando a classe 4 positiva (Área = 0,9997)



Fazendo-se a análise das figuras 10, 11, 12, 13 e 14, percebe-se uma grande melhora no classificador, de forma que o erro foi reduzido para 4,2% e a taxa de verdadeiro positivo para todas as classes ficou, aproximadamente, maior ou igual a 90%. É possível se notar a eficiência do classificador, também, a partir das curvas ROCs as quais, pelo valor de suas áreas, mostraram-se muito superiores se comparadas com as ROCs da SVM Linear.

O estudo acima foi realizado da mesma forma para outras duas situações diferentes, nas quais o número de partições foi alterado para cinco e para vinte. As variações para os valores dos erros foram bem pequenas, em torno de 0,5% à 1% para mais ou para menos.

Outra situação hipotética analisada foi a alteração do conjunto de dados através da replicação de valores das classes 2, 3 e 4, mantendo o número de partições na validação cruzada igual a dez. Para esta situação, o valor do erro para a SVM Linear caiu para 16,9% e para a SVM Quadrática caiu para 1,8%. Dessa forma, percebe-se uma melhora dos resultados quando se tem uma quantidade mais próxima de valores em cada classe. Esse resultado pode ser visto através das matrizes de confusão, mostradas nas figuras 15 e 16, para cada uma dessas situações.

Figura 15: Matriz de confusão para SVM Linear com dados replicados.

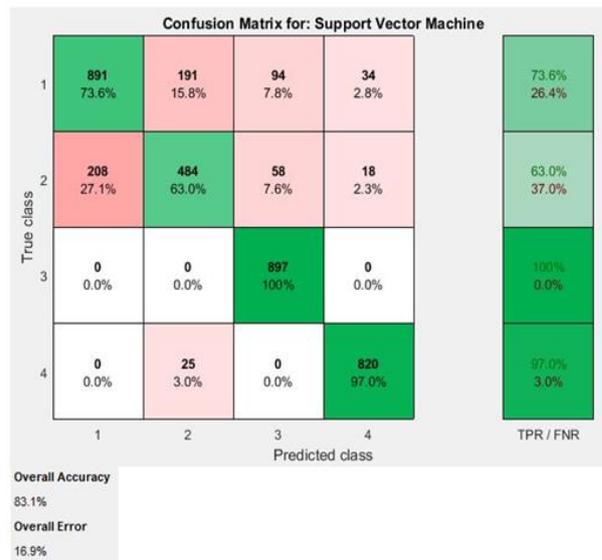
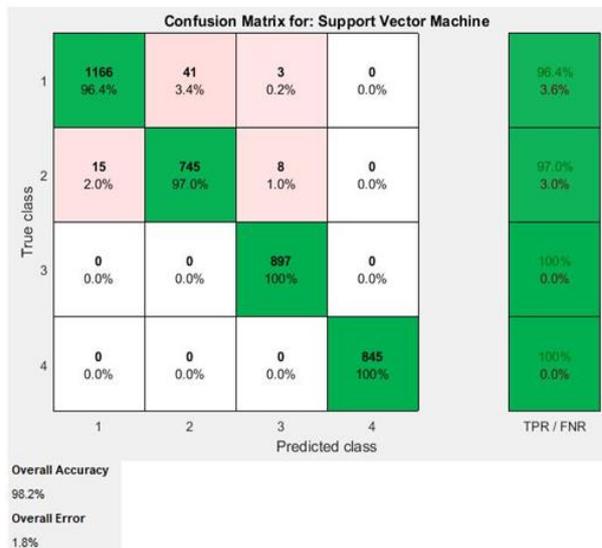


Figura 16: Matriz de confusão para SVM Quadrática com dados replicados.



### 3.2. Método MLP

Uma vez que a literatura não define uma “quantidade padrão” de elementos nas camadas ocultas, tampouco há um consenso entre os pesquisadores nos fóruns e artigos consultados, a equipe optou por variar a quantidade de elementos da camada oculta e comparar seus resultados. Considerando que o banco de dados utilizado confere um volume pequeno de dados, com baixa complexidade de classificação, a equipe adotou redes de 2, 5 e 10 elementos na camada oculta, comparando seus resultados em termos de eficiência, e posteriormente comparando os resultados obtidos com os resultados obtidos através do método de Máquina de Vetores de Suporte (SVM).

Para fins de análise dos resultados obtidos com a MLP, foram considerados os valores médios de eficiência para cada configuração (2, 5 e 10 elementos ocultos) através da média geométrica de saídas corretas para cada classe, sendo o valor de eficiência global calculado através da média geométrica da eficiência das 4 classes, conforme a tabela 2.

Tabela 2. Eficiência do método MLP para cada classe em análise.

Qtd. Elementos Camada Oculta	Eficiência Classe 1	Eficiência Classe 2	Eficiência Classe 3	Eficiência Classe 4	Eficiência Global
2	80.50%	65.06%	NaN	60.85%	68.30%
5	87.34%	71.98%	58.58%	72.20%	71.81%
10	88.76%	76.86%	84.50%	71.55%	80.14%

As classes presentes no banco de dados escolhidos foram codificadas como sendo: Unacceptable = classe 1; Acceptable = classe 2; Good = classe 3; Very Good = classe 4.

Pôde-se observar um aumento na eficiência global da rede a partir do aumento do número de elementos ocultos da rede, o que se entende como um comportamento esperado, uma vez que se aplicou um aumento da capacidade de processamento da rede.

Observa-se também, no entanto, que apesar da melhoria de eficiência decorrida do aumento do número de elementos ocultos na rede, os valores globais de eficiência não apresentam resultados satisfatórios (máximo de 80,14%), quando comparados com aplicações de Redes Neurais MLP consideradas satisfatórias.

Outro dado importante na análise são as eficiências por classe, obtidas através de média geométrica das eficiências obtidas em cada iteração. A tabela 2 mostra um valor de eficiência nulo para a classe 3 na rede com 2 elementos na camada oculta. O que significa que nesta configuração não foram detectados com êxito itens da classe 3, enquanto que com 10 elementos na camada oculta, o valor obtido foi de 84,5%.

Vale ressaltar que apesar do valor de eficiência de 84,5% obtido para a classe 3 na rede com 10 elementos ocultos ser superior à eficiência obtida para a classe 2, de 76,86%, os elementos da classe 3 apresentaram valores nulos em diversas rotinas de treinamento, nas 3 configurações possíveis da rede, conforme mostrado nas tabelas 3 e 4.

Tabela 3. Eficiência por treino para rede com 2 elementos ocultos.

Treino	Classe 1	Classe 2	Classe 3	Classe 4
1	81.9%	71.8%	NaN	NaN
2	79.7%	56.8%	NaN	NaN
3	80.0%	69.3%	NaN	62.5%
4	80.9%	62.0%	NaN	NaN
5	80.4%	68.6%	NaN	61.2%
6	80.2%	71.0%	NaN	58.9%
7	80.2%	57.7%	NaN	NaN
8	79.2%	58.0%	NaN	NaN
9	81.9%	71.6%	NaN	NaN
10	80.6%	66.4%	NaN	NaN

Tabela 4. Eficiência por treino para rede com 10 elementos ocultos.

Treino	Classe 1	Classe 2	Classe 3	Classe 4
1	98.9%	97.5%	82.6%	80.5%
2	81.6%	69.9%	NaN	NaN
3	81.0%	66.8%	NaN	NaN
4	81.7%	68.7%	NaN	NaN
5	99.3%	89.4%	90.5%	90.3%
6	80.5%	67.2%	NaN	NaN
7	96.7%	89.5%	NaN	59.6%
8	98.1%	84.0%	80.7%	NaN
9	81.5%	67.9%	NaN	NaN
10	92.0%	75.0%	NaN	60.5%

Foi constatado que a concentração das amostras do banco de dados nas classes 1 e 2, conforme figura 17, contribui para a ineficiência das redes. Essa hipótese foi comprovada gerando um banco de dados auxiliar, com mais amostras das classes 2, 3 e 4, ocasionando um ganho significativo na eficiência global da rede e das classificações para as classes 3 e 4, conforme exibido na tabela 5.

Figura 17 - Percentual de amostras por classe para o banco de dados original.

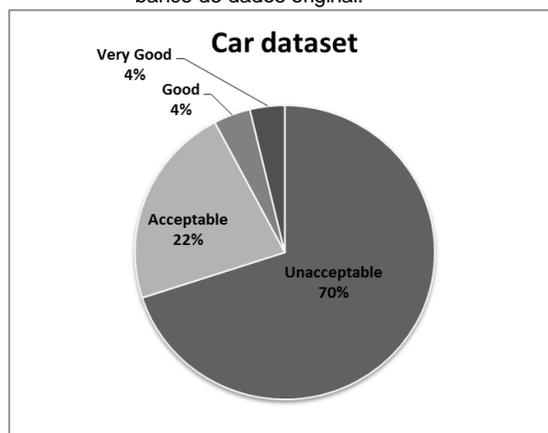


Tabela 5. Resultados de eficiência após utilização da base de dados estendida.

Qtd. Elementos Camada Oculta	Eficiência Classe 1	Eficiência Classe 2	Eficiência Classe 3	Eficiência Classe 4	Eficiência Global
2	70.63%	67.17%	86.14%	84.55%	76.67%
5	86.48%	74.24%	87.30%	91.15%	84.55%
10	97.50%	98.20%	97.90%	98.40%	94.20%

Apesar do ganho em eficiência obtido com a base de dados estendida, para fins de comparação com a SVM, serão utilizados os resultados obtidos através da classificação da base de dados original.

## 4. CONCLUSÃO

Com base nos dados coletados neste trabalho, é possível observar que a Máquina de Vetores de Suporte Quadrática foi o método com melhor resultado dentre os modelos propostos.

A SVM Quadrática apresentou resultado superior à linear, em função da característica de não-linearidade dos dados analisados. Em muitos casos, classes diferentes apresentavam mais de um atributo em comum, o que dificulta o tratamento de dados por meio de classificadores lineares.

As Redes Neurais Artificiais MLP não apresentaram desempenho satisfatório em função da diferença de quantidade de amostra de cada classe. Hipótese que foi comprovada utilizando uma base de dados auxiliar com mais amostras das classes 2, 3 e 4, que, no entanto, não teve seus resultados utilizados para termos de comparação com o SVM.

Desta forma, conclui-se que para conjuntos de dados com esta natureza, de concentração de amostras em classes específicas e quantidade restrita, o uso de Máquinas de Vetores de Suporte pode apresentar desempenho superior a uma Rede Neural MLP, uma vez que atua através da separação de cada classe das demais de maneira individual.

Já as Redes Neurais MLP se mostraram mais eficientes quando haviam mais amostras de todas as classes, de forma que os atributos comuns às classes maiores (1 -70% e 2 22%) e menores (3 e 4 - 4% cada) não induziram a erros de classificação por terem sido utilizados em quantidades significativas durante as rotinas de treino e teste.

## 5. AGRADECIMENTOS

Agradeço ao SENAI CIMATEC pela oportunidade que me foi dada em poder aprender e

aprimorar meus conhecimentos na área de Automação em Sistemas Elétricos de Potência. Agradeço ao coordenador do curso, Professor Oberdan Pinheiro, o qual sempre esteve disponível para nos auxiliar. Aos Professores Frederico Cesário, Felipe Melchert, Andrei Coelho, Sanval Hebert e todos os outros que fizeram parte dessa jornada.

Gostaria de agradecer aos meus colegas com quem dividi conhecimento e experiência durante esse período, dentre eles, André Duarte, Cléssio Arruda e Murilo Souza, peças fundamentais e de grande importância para a realização desse trabalho.

## REFERÊNCIAS

- [1] LORENA, A. C., DE CARVAHO, A. C. P. L. **Introdução às Máquinas de Vetores de Suporte**. Disponível em: <[http://conteudo.icmc.usp.br/CMS/Arquivos/argui-vos\\_enviados/BIBLIOTECA\\_113\\_RT\\_192.pdf](http://conteudo.icmc.usp.br/CMS/Arquivos/argui-vos_enviados/BIBLIOTECA_113_RT_192.pdf)> Acesso em: 20 de Outubro de 2018.
- [2] MENDOZA, Leonardo Alfredo, **Redes Neurais e Máquinas de Vetores de Suporte no reconhecimento de locutor usando coeficientes MFC e características do sinal global**, 2009. 129f. Dissertação de Mestrado – Universidade Federal de Fluminense, Niterói, 2009.
- [3] PORWAL, S., SAKAR, D. **Car Evaluation**. Disponível em: <[http://www.iitp.ac.in/~arijit/dokuwiki/lib/exe/fetch.php?media=courses:2017:cs551:13\\_report.pdf](http://www.iitp.ac.in/~arijit/dokuwiki/lib/exe/fetch.php?media=courses:2017:cs551:13_report.pdf)> Acesso em: 20 de Outubro de 2018.
- [4] M. Bohanec and V. Rajkovic: **Knowledge acquisition and explanation for multi-attribute decision making**. Disponível em: <https://archive.ics.uci.edu/ml/datasets/Car+Evaluation>
- [5] BOHANEC, M., ZUPAN, B., **Car Evaluation Data Set**. Disponível em: <http://archive.ics.uci.edu/ml/datasets/Car+Evaluation> Acesso em: 24 de Outubro de 2018.
- [6] DA SILVA, I. N., Spatti, D. H., Flauzino, R. A., Liboni, L. H. B., Alves, S. F. dos R., **Artificial Neural Networks. A Practical Course**, Ed. 2017.
- [7] HAYKIN, S. **Neural Networks and Learning Machines**, 3rd Ed.
- [8] COLOBERT, R., BENGIO, S., **Links between Perceptrons, MLPs and SVMs**. Disponível em: [https://ronan.collobert.com/pub/matos/2004\\_link\\_s\\_icml.pdf](https://ronan.collobert.com/pub/matos/2004_link_s_icml.pdf) Acesso em: 29 de Setembro de 2018.
- [9] JUNIOR, E. A. T. B., Filho, E. F. S., JUNIOR, A. C. L. F., **CLASSIFICADORES DE GÊNEROS**

**MUSICAIS USANDO MÁQUINAS DE VETOR DE SUPORTE E REDES NEURAIS.** Disponível em:

<http://www.swge.inf.br/CBA2014/anais/PDF/1569935375.pdf> Acesso em: 29 de Outubro de 2018.

[10] PASSOS, U. R. C., MATIAS, I. de O., ANDRADE, M., PASSOS, C. do E. S. O., **UM ESTUDO COMPARATIVO ENTRE TÉCNICAS DE INTELIGÊNCIA COMPUTACIONAL PARA O RECONHECIMENTO ÓTICO DE CARACTERES MANUSCRITOS.** Disponível em:

<http://cdsid.org.br/sbpo2015/wp-content/uploads/2015/08/143056.pdf> Acesso em: 29 de Outubro de 2018.

[11] **INTRODUÇÃO A REDES NEURAIS ARTIFICIAIS.** Disponível em:

[http://www.gsigma.ufsc.br/~popov/aulas/rna/uteis/RNA\\_material\\_apoio.pdf](http://www.gsigma.ufsc.br/~popov/aulas/rna/uteis/RNA_material_apoio.pdf) Acesso em: 28 de Outubro de 2018.

[12] MANZAN, J. R. G., NOMURA, S., YAMANAKA, K., **A MELHORIA NO DESEMPENHO DE MLP COM O USO DE NOVOS VETORES ALVO.** Disponível em: [http://abricom.org.br/wp-content/uploads/2016/03/st\\_23.4.pdf](http://abricom.org.br/wp-content/uploads/2016/03/st_23.4.pdf) Acesso em: 29 de Setembro de 2018.

[13] AFFONSO, E. T. F., SILVA, A. M., SILVA, M. P., RODRIGUES, T. M. D., MOITA, G. F., **USO REDES NEURAIS MULTILAYER PERCEPTRON (MLP) EM SISTEMA DE BLOQUEIO DE WEBSITES BASEADO EM CONTEÚDO.** Disponível em:

<http://www.cimec.org.ar/ojs/index.php/mc/article/viewFile/3654/3567> Acesso em: 29 de Setembro de 2018.