

UNIVERSIDADE FEDERAL DO PARANÁ

OLÍCIO DA SILVA JUNIOR

MODELAGEM DE COMPONENTES DO SISTEMA DE ARREFECIMENTO DE  
MOTORES A COMBUSTÃO BASEADA EM MÁQUINA DE VETORES DE  
SOPORTE

CURITIBA

2018

OLICIO DA SILVA JUNIOR

MODELAGEM DE COMPONENTES DO SISTEMA DE ARREFECIMENTO DE  
MOTORES A COMBUSTÃO BASEADO EM MÁQUINA DE VETORES DE  
SUPORTE

Documento de dissertação de mestrado ao Programa de Pós-Graduação em Engenharia Elétrica – PPGEE, Setor de Tecnologia da Universidade Federal do Paraná, como requisito à obtenção do grau de Mestre em Engenharia Elétrica, Departamento de Engenharia Elétrica (DELT), da Universidade Federal do Paraná.

Área de concentração Sistemas Eletrônicos

Orientador: Prof. Dr. Leandro dos Santos Coelho

CURITIBA

2018

Catálogo na Fonte: Sistema de Bibliotecas, UFPR  
Biblioteca de Ciência e Tecnologia

S856m Silva Junior, Olicio da

Modelagem de componentes do sistema de arrefecimento de motores a combustão baseado em máquina de vetores de suporte [recurso eletrônico] / Olicio da Silva Junior – Curitiba, 2018.

Dissertação (mestrado) - Pós-Graduação em Engenharia Elétrica – PPGE, Setor de Tecnologia, Universidade Federal do Paraná. Orientador: Prof. Dr. Leandro dos Santos Coelho

1. Automóveis - motores. 2. Sistemas eletrônicos. I. Universidade Federal do Paraná. II. Coelho, Leandro dos Santos. IV. Título.

CDD 621.372

Bibliotecária: Vilma Machado CRB9/1563

## **TERMO DE APROVAÇÃO**

OLICIO DA SILVA JUNIOR

### **MODELAGEM DE COMPONENTES DO SISTEMA DE ARREFECIMENTO DE MOTORES A COMBUSTÃO BASEADO EM MÁQUINA DE VETORES DE SUPORTE**

Documento de dissertação de mestrado ao Programa de Pós-Graduação em Engenharia Elétrica – PPGEE da Universidade Federal do Paraná, como requisito à obtenção do grau de Mestre em Engenharia Elétrica, Departamento de Engenharia Elétrica (DELT), da Universidade Federal do Paraná.

Orientador:

Prof. Dr. Leandro dos Santos Coelho  
Departamento de Engenharia Elétrica, UFPR

Banca examinadora:

Prof. Dr. Gideon Villar Leandro  
Departamento de Engenharia Elétrica, UFPR

Prof. Dr. Anderson Luis Szejka  
Pós-Graduação em Engenharia de Produção e Sistemas (PPGEPS), PUCPR

Curitiba, 12 de Dezembro de 2018.



MINISTÉRIO DA EDUCAÇÃO  
SETOR TECNOLOGIA  
UNIVERSIDADE FEDERAL DO PARANÁ  
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO ENGENHARIA  
ELÉTRICA

## TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ENGENHARIA ELÉTRICA da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de OLÍCIO DA SILVA JUNIOR intitulada: **Modelagem de Componentes do Sistema de Arrefecimento de Motores a Combustão Baseado em Máquina de Vetores de Suporte**, após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rto de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 12 de Dezembro de 2018.

LEANDRO DOS SANTOS COELHO  
Presidente da Banca Examinadora (UFPR)

ANDERSON LUIS SZEJKA  
Avaliador Externo (PUC/PR)

GIDEON VILLAR LEANDRO  
Avaliador Interno (UFPR)

A minha esposa Maria Elena que nos momentos mais difíceis sempre me encorajou a não desistir, aos meus filhos Pedro e Lucas que são minha motivação. Aos meus pais Florinda e Olicio, meus exemplos e força sempre! Aos meus irmãos Juliane, Ronaldo e Rodrigo.

## **AGRADECIMENTOS**

Ao meu maior incentivador, o amigo Carlos Eduardo Klein.

Ao meu professor orientador Leandro dos Santos Coelho e aos professores Viviana Coco Mariani, Alexandre Rasi Aoki, Daniel Pipa e Roberto Zanetti Freire por seus exemplos como professores e excelentes profissionais.

Aos professores das bancas de qualificação e defesa, Gideon Villar Leandro, Eduardo Parente e Anderson Luis Szejka, obrigado pela dedicação nas correções e pelas valiosas dicas.

Aos amigos do curso de mestrado Eduardo Ribeiro, Edson Gruska, Thaisa Oda, Cristiano Osinski e Victor Henrique Alves Ribeiro.

Aos amigos Roberto Florindo, Mario Bittencourt, Samuel Giatti, Iron Tessaro, pelo apoio. Teodoro Bora, Erik Akiyama, Marcos Limas e Igor Campos, pela ajuda com o estudo do sistema e na aquisição dos dados.

A Bruna Krasota Matos e Rodrigo da Silva pela ajuda com as revisões.

A VOLVO do Brasil Veículos Limitada por incentivar e permitir que seus funcionários possam dedicar tempo em atividades de mestrado e doutorado, agradeço aos amigos da Engenharia de Sistemas de Controle pela paciência na convivência durante o longo período deste mestrado.

A Deus, graças ao bom Pai sempre!

Não existe um problema que não ofereça uma dádiva para você. Você procura os problemas porque precisa das dádivas por eles oferecidas.

(Richard Bach, 1977)

## RESUMO

A área de desenvolvimento de veículos automotivos teve ganhos com a introdução dos motores eletrônicos. A indústria automotiva tem que lidar com limites, desafiadores de operação dos motores a combustão. Limites impostos pela legislação de controle de emissões de poluentes na atmosfera, além de uma demanda crescente por economia de combustível imposta pelos dos usuários de caminhões com motores Diesel. Conseqüentemente a indústria tem buscado soluções que consigam aliar estas duas condições para poder vender veículos. O sistema de arrefecimento é fundamental para que um motor a combustão funcione, uma vez que a temperatura de operação do motor está diretamente ligada à queima de combustível, este possui um comportamento de controle complexo apesar do mesmo ser gerenciado eletronicamente. Nesse contexto, essa dissertação propôs-se a obter um modelo para a rotação do ventilador e para a temperatura do fluido do sistema de arrefecimento com o auxílio de métodos de Aprendizado de máquina, a partir de dados reais de medições realizadas em caminhão. Entre os algoritmos de Aprendizado de máquina avaliados estão as redes neurais artificiais, do tipo perceptron multicamadas, que se caracterizam por serem modelos caixa-preta para modelagem não-linear. Outra técnica de Aprendizado de máquina adotada foi a máquina de vetores de suporte para regressão e sua variação aplicada com o método dos mínimos quadrados por batelada com inferência estatística, fundamentada na teoria do aprendizado estatístico. O objetivo de tais técnicas é maximizar a capacidade de generalização em problemas de regressão. Os resultados obtidos, para a modelagem dos dois componentes do sistema de arrefecimento, mostraram que a máquina de vetores de suporte para regressão (regressão por vetores de suporte), quando utilizada em combinação com modelo linear estimado pelo algoritmo de mínimos quadrados em batelada com inferência estatística, obteve um desempenho superior ao das redes neurais perceptron multicamadas e também da máquina de vetores de suporte para regressão aplicada de forma isolada (sem a combinação com o modelo linear). Os modelos apresentaram desempenho acima de 80% para o sinal de rotação e acima de 60% para o modelo da temperatura, indicando que muito ainda pode ser realizado para a obtenção de modelos robustos e úteis para a indústria automotiva.

**Palavras-chave:** Motor a combustão; Ventilador; Sistema de arrefecimento; Máquina de vetores de suporte; Rede neural perceptron multicamadas; Regressão não-linear.

## ABSTRACT

The automotive vehicle development area had gained with the introduction of the electronic to engines. The automotive industry has to deal with challenging limits for the combustion engine operation, due to the legislation that controls the pollutant emission from those engines to the atmosphere and also due to a growing demand from the Diesel truck drivers for fuel economy. Therefore, the automotive industry has been searched for solutions which could solve both constrains in order to be able to sell vehicles. The cooling system is fundamental for the proper engine operation, once the engine operating temperature is directly linked to the fuel burning. That system has a complex control behavior besides being electronic managed. In this context. this dissertation aimed by means of the Machine learning algorithms, employing real data from measurements done in the truck, to estimate the fan speed and also the coolant temperature from that system. Among the assessed algorithms, are the perceptron multilayer artificial neural network, which are classified as black-box models for non-linear modeling. Another employed Machine learning technique was the support vector machine applied to regression and its variation with least squares method in batch processing and statistical inference, this one grounded on the statistical learning theory. Machine learning algorithms aim is to maximize the generalization capacity within regression problems. The achieved results, for modeling the two cooling system components, showed that support vector machine applied to regression (regression by means of support vectors), when working with combination to the batch least squares method for linear model generation in batch processing and statistical inference, has achieved the higher performance than the perceptron multilayer artificial neural networks and also than the support vector machine applied to regression when isolated applied (without combination with the linear model). The models showed performance superior to 80% for the fan speed and 60% superior the cooling temperature model, these results indicate that much more could be performed to obtain robust and useful models to the automotive industry.

**Key-words:** *Combustion engine; Fan; Cooling system; Support Vector Machine; Multilayer perceptron neural network; Non-linear regression.*

## LISTA DE FIGURAS

Figura 1.1 – Sistema de arrefecimento de carros de passeio .....	15
Figura 1.2 – Etapas de trabalho da dissertação .....	20
Figura 2.1 – Diagrama do circuito de arrefecimento fechado (a), aberto (b) .....	25
Figura 2.2 – Ciclos de acoplamento e desacoplamento do ventilador e histereses ..	27
Figura 2.3 – Embreagem Visco® para ventiladores automotivos.....	28
Figura 2.4 – Diagrama do ventilador com controle eletrônico .....	29
Figura 3.1 – Classes de modelos e sua confiabilidade, funcionamento x dados .....	35
Figura 3.2 – Modelo de sistemas lineares: superposição, aditividade e homogeneidade.....	36
Figura 3.3 – Fases de aprendizado supervisionado .....	39
Figura 3.4 – Par de distribuições de dados bidimensional (dimensão VC) .....	46
Figura 3.5 – Espaços de entrada e de características .....	46
Figura 3.6 – Fronteira de decisão de separação de classes .....	47
Figura 3.7 – Separação ótima entre duas classes (SV em cinza) .....	47
Figura 3.8 – O mapeamento de espaços na SVM.....	48
Figura 3.9 – Hiperplanos de alta dimensão para separação de classes .....	49
Figura 3.10 – Exemplo de estimação de função Sinc x com aproximação $\epsilon$ -SVR ....	50
Figura 3.11 – Vetores de suporte no $\epsilon$ -SVR.....	50
Figura 3.12 – Separação linear de hiperplanos.....	52
Figura 3.13 – Margens soft, funções de aproximação $\epsilon$ -SVR (esquerdo) e perda $\epsilon$ - insensível (direito) .....	53
Figura 3.14 – Função de perda $\epsilon$ -insensível com variáveis de folga $\xi_i, \xi_i^*$ .....	54
Figura 3.15 – Mapeamento não-linear $\phi_i$ do espaço de entrada para o de características .....	58
Figura 3.16 – Neurônio artificial de McCulloch-Pitts.....	64
Figura 3.17 – Grafo de fluxo de sinal de um neurônio.....	66
Figura 3.18 – Rede alimentada adiante, uma camada de neurônios .....	68
Figura 3.19 – Rede com uma camada oculta e uma de saída .....	69
Figura 3.20 – Rede neural recorrente .....	70
Figura 3.21 – Diagrama de blocos de uma ANN, um neurônio na camada de saída e erro.....	71

Figura 4.1 – Diagrama de trabalho dos dados adquiridos .....	75
Figura 4.2 – Diagrama MISO de entradas e saída dos modelos.....	76
Figura 5.1 - Rotação do ventilador com LSM .....	89
Figura 5.2 – Validação rotação do ventilador para LSM.....	90
Figura 5.3 – Temperatura do fluido com LSM .....	90
Figura 5.4 – Validação temperatura do fluido para LSM .....	91
Figura 5.5 – Rotação do ventilador com StatLSSVM .....	92
Figura 5.6 – Validação rotação do ventilador para StatLSSVM .....	92
Figura 5.7 – Temperatura do fluido com StatLSSVM .....	93
Figura 5.8 – Validação temperatura do fluido para StatLSSVM .....	93
Figura 5.9 – Rotação do ventilador com SVR .....	94
Figura 5.10 – Validação rotação do ventilador para SVR.....	94
Figura 5.11 – Temperatura do fluido com SVR .....	95
Figura 5.12 – Validação temperatura do fluido para SVR .....	95
Figura 5.13 – Rotação do ventilador com ANN .....	96
Figura 5.14 – Validação rotação do ventilador para .....	97
Figura 5.15 – Temperatura do fluido com ANN .....	97
Figura 5.16 – Validação temperatura do fluido para ANN .....	98

## LISTA DE QUADROS

Quadro 3.1 – Técnicas de aprendizado de máquina .....	39
Quadro 3.2 – Funções <i>kernel</i> aplicáveis a SVM.....	58

## LISTA DE TABELAS

Tabela 4.1 – Variáveis adquiridas durante as medições .....	76
Tabela 4.2 – Correlação das entradas/saída do modelo LSM para rotação do ventilador.....	77
Tabela 4.3 – Correlação das entradas/saída do modelo LSM para temperatura do fluido.....	78
Tabela 4.4 – Medições por trajetos x amostras.....	79
Tabela 4.5 – Proporção dados treinamento e validação para rotação do ventilador.	80
Tabela 4.6 – Proporção dados treinamento e validação para temperatura do fluido de arrefecimento .....	80
Tabela 4.7 – <i>Kernels</i> para a StatLSSVM no caso da rotação ventilador.....	81
Tabela 4.8 – <i>Kernels</i> para a StatLSSVM no caso da temperatura do fluido .....	82
Tabela 4.9 – Parâmetros de ajuste para a StatLSSVM.....	82
Tabela 4.10 – Função <i>kernel</i> para a velocidade do ventilador .....	83
Tabela 4.11 – Função <i>kernel</i> para a temperatura do fluido.....	83
Tabela 4.12 – Funções de ativação para ANN – rotação do ventilador (20 neurônios) .....	83
Tabela 4.13 – Funções de ativação para ANN – temperatura do fluido (50 neurônios) .....	84
Tabela 4.14 – Número de neurônios para ANN – rotação do ventilador.....	84
Tabela 4.15 – Número de neurônios para ANN – temperatura do fluido.....	84
Tabela 5.1 – Treinamento e validação das técnicas para a rotação do ventilador ....	85
Tabela 5.2 – Treinamento e validação das técnicas para a temperatura do fluido....	85
Tabela 5.3 – Comparativos da validação para rotação do ventilador.....	86
Tabela 5.4 – Comparativos da validação para temperatura do fluido .....	86
Tabela 5.5 – Comparativos de treinamento e validação para rotação do ventilador .	87
Tabela 5.6 – Comparativos de treinamento e validação para temperatura do fluido.	88

## LISTA DE ABREVIATURAS E SIGLAS

AM	Aprendizado de máquina (do inglês <i>Machine Learning</i> )
ANN	Rede Neural Artificial (do inglês <i>Artificial Neural Network</i> )
ARX	Auto regressivo com entradas exógenas (do inglês <i>AutoRegressive with eXogenous inputs</i> )
BIBO	Entrada limitada, saída limitada (do inglês <i>bounded-input, bounded-output</i> )
CFD	Ferramenta computacional para dinâmica dos fluidos (do inglês <i>Computational Fluid Dynamics, CFD</i> )
CO	Monóxido de carbono
CB	Modelo caixa branca
CC	Modelo caixa cinza
CP	Modelo caixa preta
CONAMA	Conselho Nacional do Meio Ambiente
ECU	Unidade de Controle Eletrônico (do inglês <i>Electronic Control Unit</i> )
EGR	Válvula de recirculação do gás de escape (do inglês <i>Exhaust Gas Recirculation</i> )
HC	Hidrocarboneto
IBAMA	Instituto Brasileiro do Meio Ambiente
KKT	Condições de Karush-Kuhn-Tucker
LS	Mínimos quadrados (do inglês <i>Least Square</i> )
LSM	Método dos Mínimos Quadrados (do inglês <i>Least Square Method</i> )
LSSVM	SVM baseada em mínimos quadrados (do inglês <i>Least Square Support Vector Machine</i> )
LTI	Sinal linear invariante no tempo (do inglês <i>Linear Time-Invariant</i> )
MISO	Múltiplas entradas-saída única (do inglês <i>Multiple Inputs-Single Output</i> )
MLP	Rede neural Perceptron multicamadas (do inglês <i>Multi-layer Perceptron</i> )
MSE	Erro Quadrado Médio (do inglês <i>Mean Squared Error</i> )
NARX	Auto regressivo não-linear com entradas exógenas (do inglês <i>Nonlinear AutoRegressive with eXogenous inputs</i> )

NEDC	novo ciclo europeu para medição de emissões em motores a combustão interna (do inglês <i>New European Driving Cycle</i> )
NOx	Óxido de nitrogênio (Enxofre)
PCA	Análise de Componentes Principais (do inglês <i>Principal Component Analysis</i> )
PM	Material Particulado (do inglês <i>Particulated Matter</i> )
PMI	Ponto Morto Inferior do cilindro do motor
PMS	Ponto Morto Superior do cilindro do motor
PROCONVE	Programa de controle de emissões veiculares
PROMOT	Programa de Controle de emissões veiculares por Motociclos e Veículos Similares
R <sup>2</sup>	Coefficiente de correlação múltipla
RBF	Função de Base Radial (do inglês <i>Radial Basis Function</i> )
RSS	Soma residual dos quadrados (do inglês <i>Residual Sum of Squares</i> )
SMO	Otimização mínima sequencial (do inglês <i>Sequential Minimal Optimization</i> )
SS	Soma dos quadrados devido regressão (do inglês <i>Sum of Squares to regression</i> )
StatLSSVM	LSSVM com inferência estatística (do inglês <i>Statistical LSSVM</i> )
SVM	Máquina de Vetores de Suporte (do inglês <i>Support Vector Machine</i> )
SVR	Vetores de Suporte para Regressão (do inglês <i>Support Vector Regression</i> )
SV	Vetor de Suporte (do inglês <i>Support Vector</i> )
TAE	Teoria do Aprendizado Estatístico (do inglês <i>Statistical Learning Theory</i> )
TSS	Soma total dos quadrados (do inglês <i>Total Sum of Squares</i> )
VC	Dimensão de Vapink-Chervonenkis
Wi-fi	Fidelidade sem fio (do inglês <i>Wireless Fidelity</i> )

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>13</b>
1.1	OBJETIVOS .....	17
1.1.1	Objetivo geral .....	17
1.1.2	Objetivos específicos.....	17
1.1.3	Contribuições.....	18
1.2	METODOLOGIA.....	19
1.2.1	Etapas da pesquisa .....	20
1.2.2	Organização da dissertação .....	22
<b>2</b>	<b>FUNDAMENTOS BÁSICOS E REVISÃO BIBLIOGRAFICA.....</b>	<b>23</b>
2.1	O MOTOR A COMBUSTÃO INTERNA .....	23
2.2	CIRCUITO DE ARREFECIMENTO .....	24
2.3	CONTROLE DE ACIONAMENTO DO VENTILADOR.....	26
2.4	REVISÃO BIBLIOGRÁFICA .....	30
<b>3</b>	<b>FUNDAMENTOS DE MODELAGEM .....</b>	<b>33</b>
3.1	MODELAGEM DE UM SISTEMA .....	33
3.1.1	Classificação de modelos .....	34
3.1.2	Sistemas lineares e não lineares.....	36
3.1.3	Tipos de aprendizado .....	38
3.2	MÉTODO DE REGRESSÃO LINEAR .....	40
3.2.1	Método dos mínimos quadrados .....	41
3.2.2	Mínimos quadrados recursivo.....	43
3.3	MÁQUINA DE VETOR DE SUPORTE .....	44
3.3.1	Dimensão VC .....	44
3.3.2	Espaço de características.....	46
3.3.3	A margem de separação .....	47
3.3.4	Hiperplano ótimo.....	48
3.3.5	Vetores de suporte de regressão .....	49
3.3.6	O problema de regressão .....	51
3.3.7	Funções <i>kernel</i> .....	57
3.3.8	Método SVM por mínimos quadrados .....	61
3.4	REDES NEURAIS ARTIFICIAIS.....	63
3.4.1	Neurônio artificial.....	63

3.4.2	Funções de ativação.....	65
3.4.3	Grafos orientados .....	66
3.4.4	Arquiteturas de rede .....	67
3.4.5	Processos de aprendizado .....	70
3.5	CRITÉRIOS DE DESEMPENHO.....	72
<b>4</b>	<b>PREPARAÇÃO DE DADOS E AJUSTES DAS TÉCNICAS.....</b>	<b>74</b>
4.1	AQUISIÇÃO DOS DADOS .....	74
4.2	ESCOLHA DOS DADOS DE ENTRADA E SAÍDA DO MODELO .....	75
4.2.1	Correlação dos dados dos modelos .....	76
4.3	NORMALIZAÇÃO DOS DADOS.....	78
4.4	ESCOLHA DO TRAJETO REALIZADO PELO VEÍCULO DE REFERÊNCIA 78	
4.5	DIVISÃO DOS DADOS DE TREINAMENTO E VALIDAÇÃO.....	80
4.6	AJUSTES ESPECÍFICOS DOS PARÂMETROS DE CADA TÉCNICA .....	81
4.6.1	Função <i>kernel</i> para StatLSSVM .....	81
4.6.2	Parâmetros de ajuste para a StatLSSVM.....	82
4.6.3	Função <i>kernel</i> para SVR .....	82
4.6.4	Função de ativação para ANN.....	83
4.6.5	Número de neurônios para ANN .....	84
<b>5</b>	<b>RESULTADOS .....</b>	<b>85</b>
5.1	TESTE DE ESTABILIDADE DOS MODELOS.....	85
5.2	TREINAMENTO E VALIDAÇÃO.....	87
5.3	REPRESENTAÇÃO GRÁFICA DOS RESULTADOS.....	88
5.3.1	Modelagem LSM.....	89
5.3.2	Modelagem StatLSSVM .....	91
5.3.3	Algoritmos SVR .....	93
5.3.4	Algoritmos ANN .....	96
<b>6</b>	<b>CONCLUSÃO.....</b>	<b>99</b>
	<b>REFERÊNCIAS.....</b>	<b>101</b>

## 1 INTRODUÇÃO

Os motores a combustão interna começaram a ser utilizados em veículos no início do século XX e desde então tem passado por diversas etapas de desenvolvimento e aprimoramento em termos de desempenho, conforto e segurança. Segundo Ashok et al. (2016), as demandas rigorosas por redução de emissões de poluentes na atmosfera, redução do consumo de combustível e aumento da segurança tem justificado técnicas avançadas de controle do motor o qual tem-se tornado um sistema mecatrônico composto por eletrônica embarcada com sensores, atuadores e microprocessadores.

Veículos são lançados periodicamente para atender a diferentes tipos de aplicações, como usuários particulares, veículos coletivos, utilitários e para o transporte de cargas pesadas. Em cada uma dessas aplicações, são necessárias configurações específicas dos veículos, como por exemplo, a definição do número de eixos de rodas em veículos pesados, arranjos de composição de cargas além de motores com potências variadas (conforme Bosch<sup>1</sup>, 2004). Nos últimos anos os veículos leves ou de passeio passaram a utilizar motores acionados por combustão interna a gasolina, álcool ou diesel. Os veículos médios e pesados são em sua maioria movidos por motores Diesel, que devido à sua curva característica de torque, torna-os veículos robustos e capazes de transportar cargas grandes e pesadas.

O estreitamento das legislações de emissões para veículos a diesel, tem causado impacto tecnológico nos futuros veículos que circularão nas rodovias, tais como a Legislação Euro VI proposta em 2014 na Europa e o PROCONVE P7 (Programa de Controle da Poluição do Ar por Veículos Automotores) proposto em 2012 no Brasil.

Com o objetivo de reduzir e controlar a contaminação atmosférica e a emissão de ruído por fontes móveis (veículos automotores) o Conselho Nacional do Meio Ambiente - CONAMA criou os Programas de Controle da Poluição do Ar por Veículos Automotores: PROCONVE (automóveis, caminhões, ônibus e máquinas rodoviárias e agrícolas) e PROMOT (motocicletas e similares) fixando prazos, limites máximos de emissão e estabelecendo exigências tecnológicas para veículos automotores, nacionais e importados. CONAMA IBAMA (2017)

---

<sup>1</sup> Matriz de veículos rodoviários e Carroceria (utilitários).

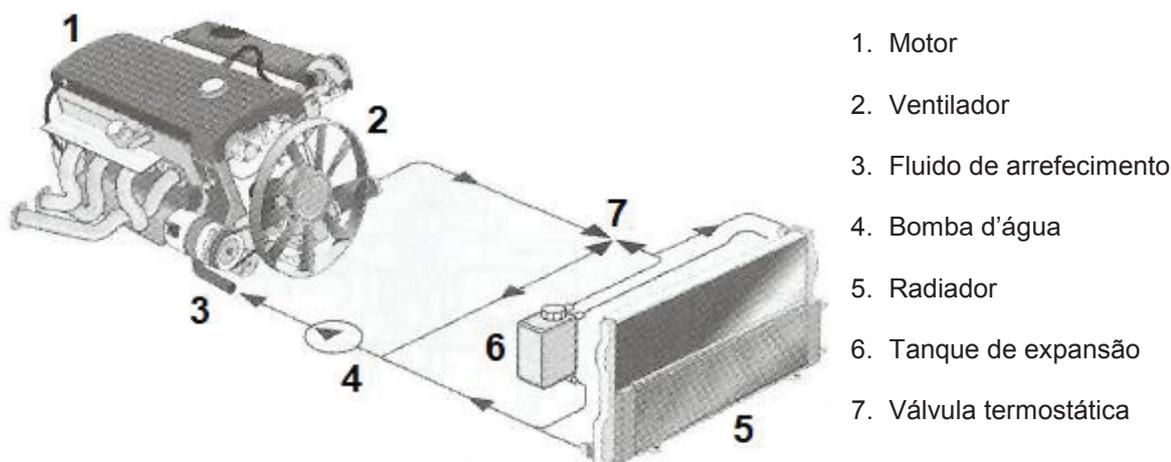
A operação de veículos a combustão capazes de atender às imposições da PROCONVE P7 tem sido possível graças à utilização de controladores eletrônicos geridos por *software* embarcado dependentes de dispositivos eletrônicos como sensores e atuadores, aumentando assim a complexidade tecnológica em termos de calibração e harmonização entre os sistemas do veículo, segundo Isermann e Haffner (2001), alguns componentes mecânicos clássicos passaram a ser mecatrônicos, como por exemplo o controle do comando de válvulas (*camshaft*). Em contrapartida, nem todos os sensores puderam ser adicionados aos veículos de produção por serem caros e terem baixa durabilidade, assim a pressão pela redução de custos motiva que os fabricantes adotem configurações enxutas de seus veículos com número mínimo de componentes para funcionar, isso torna o controle por *software* limitado.

Um motor a combustão é composto por sistemas dinâmicos e não-lineares, como por exemplo o sistema de combustão. Nesse sistema os ciclos contínuos produzem diferentes temperaturas que se não forem controladas adequadamente causam desgastes prematuros dos componentes do motor. Logo o controle se faz necessário para proporcionar a segurança do motorista e para garantir a durabilidade do motor.

Dentre os sistemas que utilizam controle eletrônico destaca-se o de arrefecimento, responsável pelo controle da temperatura do motor. Segundo Khodabakhshian et al. (2014), este sistema é caracterizado por ser não-linear e de difícil modelagem e será objeto de estudo desta dissertação.

O controle da temperatura da água do motor é feito por meio do radiador e demais componentes conforme ilustrados na Figura 1.1.

Figura 1.1 – Sistema de arrefecimento de carros de passeio



FONTE: Adaptado de Bosch (2004)

Segundo Bosch (2004), com o passar do tempo os líquidos de arrefecimento evoluíram da água pura para uma mistura de água, anticongelante e vários inibidores de corrosão, conhecida como fluido de arrefecimento. Essa mistura eleva o ponto de ebulição permitindo que o fluido opere com uma temperatura de até 120°C e uma pressão de 1,4 Bar em carros de passeio.

O sistema também é composto por um ventilador, cujo acionamento pode ser elétrico, eletromecânico ou simplesmente mecânico. Quando um veículo está em baixa velocidade o ventilador produz uma ventilação forçada sobre o radiador de modo a manter o equilíbrio térmico do motor a combustão. Em paralelo o fluido de arrefecimento passa pelo bloco do motor ajudando no seu resfriamento. O controle da temperatura evita o sobreaquecimento do bloco metálico do motor e suas partes móveis, tais como pistões e válvulas. O ventilador e a bomba d'água são controlados pela unidade de controle eletrônico responsável pelo cálculo da demanda de calor a ser rejeitada no radiador por meio da troca de calor entre o fluido de arrefecimento e o ar que passa pelo radiador.

Segundo Heywood (1988) a transferência de calor afeta o desempenho, a eficiência e as emissões do motor. Portanto, se o motor trabalhar na temperatura ideal terá maior durabilidade e economia de combustível além de auxiliar na redução de emissões de poluentes na atmosfera, sendo assim o sistema de arrefecimento é um dos mais importantes para garantir a longevidade do motor. O ventilador do sistema consome parte da energia produzida pelo motor a qual deixa de ser transferida às rodas, afetando assim o desempenho e consumo do veículo.

Em veículos pesados um ventilador totalmente acoplado chega a consumir cerca de 100 kW podendo chegar a um décimo da potência em motores Diesel. Além disso o sistema de arrefecimento possui uma dinâmica complexa e de difícil controle, como o acionamento viscoso do ventilador e sua histerese nos períodos de acoplamento e desacoplamento (estes serão detalhados no Capítulo 2 desta dissertação).

O gerenciamento do motor é um dos mais complexos na área de software de controle automotivo. Funções como cálculo do tempo e quantidade de injeção de combustível demandam precisão e rapidez de processamento. O controle do solenoide do ventilador do sistema de arrefecimento necessita de um ajuste fino em seu comportamento, devido à sua característica de lentidão de resposta como descrito anteriormente. Nestes casos a otimização é feita pela calibração de parâmetros existentes no software, como visto em Grelet et al. (2015). Segundo Huang et al. (2010), engenheiros e cientistas dependem cada vez mais de computadores para processar uma enorme quantidade de dados (*big data*) fornecidos por sensores. O objetivo do Aprendizado de máquina AM (do inglês *Machine Learning*), é extrair conhecimento desses dados e utilizar computadores para tomar decisões complexas. Neste contexto, o AM pode ser útil em decisões como o melhor momento para injetar combustível e no aprendizado da relação de entre eficiência e geração de resíduos de combustão controlados pelas legislações de emissões de poluentes na atmosfera como mencionado por Pang et al. (2004).

O AM poderia também uma abordagem viável para aprender as situações de demanda por resfriamento do motor prevendo com eficiência a atuação do ventilador em função do atraso causado por sua dinâmica de trabalho.

Esta dissertação não aborda o sistema de arrefecimento por completo, mas propõe-se a modelar duas características importantes desse sistema: a rotação de controle do ventilador e da temperatura do fluido de arrefecimento, empregando-se identificação de sistemas por meio de técnicas de AM para a construção dos modelos.

Segundo Aguirre (2015), a modelagem matemática é a área do conhecimento que estuda maneiras de desenvolver e implementar modelos matemáticos de sistemas reais, sendo a identificação de sistemas a área do conhecimento responsável pelo estudo da técnica de modelagem matemática. Para Rabelo (2018), AM refere-se a uma máquina capaz de aprender por meio de

informações ou dados fornecidos e da utilização de algoritmos capazes de reconhecer. Segundo Kung (2014), AM é uma área de pesquisa relacionada ao projeto de algoritmos adaptativos capazes de inferir informações relevantes baseados em dados de treinamento, com isso induz as melhores regras de decisão para classificação, agrupamento ou regressão.

Nesta dissertação foram adotadas as Redes Neurais Artificiais (do inglês *Artificial Neural Network* - ANN) e a Máquina de Vetores de Suporte para regressão (do inglês, *Support Vector Machine* - SVM para regressão ou simplesmente SVR). As ANNs e os métodos baseados em SVM são dois importantes paradigmas de AM. Para Horta e Braga (2013) esses algoritmos têm em comum a capacidade de produzir um modelo matemático capaz de modelar comportamentos não-lineares e serão abordados no Capítulo 3. Também foi avaliado o Método dos Mínimos Quadrados (do inglês, *Least Mean Squares* - LSM) para o ajuste de modelos lineares.

## 1.1 OBJETIVOS

A seguir são apresentados os objetivos, geral e específicos, desta dissertação.

### 1.1.1 Objetivo geral

O objetivo geral da dissertação é gerar modelos para rotação do ventilador com acoplamento viscoso, com controle eletrônico e para a temperatura do fluido do sistema de arrefecimento do motor Diesel, utilizando técnicas de Aprendizado de Máquina.

### 1.1.2 Objetivos específicos

Esta dissertação foi elaborada visando atender aos objetivos específicos a seguir de modo a atingir o objetivo geral proposto:

- a) Estudar o sistema de arrefecimento do motor Diesel em um caminhão, observando a complexidade deste sistema e as dificuldades em realizar o seu controle, bem como os seus benefícios.
- b) Definir o que seria útil de se modelar para os futuros trabalhos de calibração do sistema de arrefecimento, qual saída do modelo matemático poderia representá-lo e na sequência definir quais sinais de entrada seriam necessários para a obtenção do modelo desejado.
- c) Coletar dados reais do sistema de arrefecimento em diferentes medições e prepará-los para serem usados na modelagem do sistema.
- d) Escolher as técnicas de aprendizado de máquina a serem abordadas nesta dissertação.
- e) Obter os modelos matemáticos gerados pelas técnicas de rede neural do tipo perceptron multicamadas e SVM para regressão, além de modelo linear ajustados por mínimos quadrados por batelada. Fazer os ajustes específicos de cada técnica e validar a eficácia de tais modelos.

### 1.1.3 Contribuições

Como contribuição desta dissertação tem-se que a modelagem da rotação do ventilador e da temperatura do fluido realizada por abordagens de Aprendizado de máquina seja precisa e relevante para um melhor entendimento das características complexas do sistema de arrefecimento. Além disso, espera-se que os modelos dos componentes de arrefecimento do motor Diesel possam fornecer subsídios para o refinamento da calibração do sistema por meio de reutilização customizada dos modelos em outros veículos.

Assim como a modelagem da temperatura do fluido baseada no comportamento do sistema pode ser vista como um sensor virtual e ser utilizada, por exemplo, como referência para a função de monitoramento de possíveis desvios no comportamento dinâmico do sistema de arrefecimento, proporcionando subsídios para a realização de manutenção preventiva do motor de modo a garantir que a temperatura do motor seja corrigida em tempo hábil. Em casos de falha do sensor a temperatura prevista poderá ser usada como modelo auxiliar na tomada

de decisão, permitindo que o veículo seja levado em segurança para a oficina evitando paradas em situações e locais de risco.

## 1.2 METODOLOGIA

Neste capítulo são apresentados detalhes da pesquisa, uma breve introdução sobre as técnicas usadas para a geração do modelo para o sistema de arrefecimento, bem como os critérios de avaliação aplicados. Algumas referências permanecem válidas ao longo dos anos, pois, referem-se à dinâmica fundamental do sistema.

O objeto de estudo é o sistema de arrefecimento com ventilador acionado por acoplamento do fluido viscoso e controle eletrônico utilizado no Motor de 13l a combustão Diesel montado em um caminhão VOLVO modelo FH calibrado para potências de trabalho de 460 e 540 hp (do inglês *horse power*) e sistema de pós tratamento dos gases de exaustão para atender a legislação brasileira PROCONVE P7, o qual será denominado de veículo de referência nesta dissertação.

A coleta de dados foi realizada durante testes de validação com o veículo de referência carregado em sua carga nominal de 14 toneladas, operando em diferentes trajetos e ciclos de trabalho (detalhes são apresentados no Capítulo 4). Os dados coletados do motor foram: rotação, torque, temperatura do motor. Para o ventilador foram coletados dados de rotação e suas respectivas demandas de acionamento.

Para geração dos modelos matemáticos utilizou-se a implementação de algoritmos de Aprendizado de máquina, estas implementadas em ambiente computacional Matlab®.

As técnicas utilizadas foram

- Redes neurais artificiais (do inglês *Artificial Neural Network* - ANN) com Perceptron multicamadas;
- Vetores de suporte de regressão (do inglês *Support Vector Regression* - SVR);

além da técnica clássica de modelagem conhecida como

- Método dos mínimos quadrados (do inglês, *Least Squares Method* - LSM).

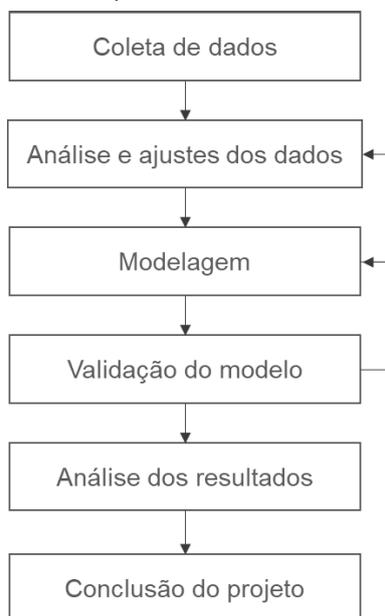
### 1.2.1 Etapas da pesquisa

Segundo Fonseca (2002) uma pesquisa pode ser classificada como qualitativa e quantitativa, como onde:

- A pesquisa qualitativa se preocupa com os aspectos da realidade que não podem ser quantificados como significados, crenças, valores e atitudes. Leva em consideração a subjetividade do pesquisador.
- A pesquisa quantitativa busca resultados que possam se quantificados, como amostras grandes o suficiente para representar uma população. Centra-se na objetividade, busca representar a realidade por meio de dados brutos recolhidos com o auxílio de instrumentos padronizados e neutros.

A pesquisa para a dissertação pode ser classificada como qualitativa com método empírico, uma vez que utiliza dados reais coletados do objeto de estudo informado. As etapas de trabalho para o desenvolvimento dessa dissertação são apresentadas na Figura 1.2.

Figura 1.2 – Etapas de trabalho da dissertação



FONTE: O autor (2018).

- a) Coleta de dados – A coleta dos dados foi feita com o veículo de referência rodando em diferentes condições denominadas ciclos de

trabalho, caracterizados por transportes de cargas realizados em rodovias planas, serras e nível do mar. Os dados adquiridos contemplam tais trajetos de modo a representar diferentes situações de demandas do sistema de arrefecimento e de acionamento do ventilador do motor;

- b) Análise e ajuste dos dados – Inicialmente os dados coletados do sistema de arrefecimento do veículo de referência foram normalizados conforme descrito no Capítulo 4.1. Durante o processo de modelagem algumas variáveis foram adicionadas e outras retiradas, sendo combinadas e testadas por tentativa e erro de modo a se escolher as de melhor representação do sistema avaliado segundo os critérios de desempenho adotados para modelagem, sendo que estes são descritos na seção 5 do Capítulo 3;
- c) Modelagem – Para efetivar a realização do treinamento dos modelos foi necessário definir a quantidade de dados adquiridos para serem adotados na fase de treinamento. Foram testadas diferentes quantidades de dados para o treinamento dos modelos, conforme abordado em Meruelo et al. (2016) e Paul e Kumar (2002). Após os testes decidiu-se em um utilizar partições de dados para o treinamento específicas de acordo com a técnica utilizada, conforme descrito no Capítulo 4.5.
- d) Validação dos modelos – Para a fase de validação foram utilizados o restante dos dados adquiridos e ainda não apresentados aos modelos durante a fase de treinamento, assim como em Karimpour et al. (2018);
- e) Critérios de desempenho – Um dos métodos mais comuns para a quantificação do erro dos estimadores é o erro quadrado médio (do inglês *Mean Squared Error* – MSE) no qual pode-se garantir a qualidade dos estimadores ou previsores. Outro critério escolhido foi coeficiente de correlação múltipla  $R^2$  (do inglês *multiple correlation coefficient*);
- f) Análise de resultados – Os modelos gerados para o sistema de arrefecimento foram avaliados pelos critérios de desempenho de modo a se comparar os métodos empregados, sendo as duas

abordagens de AM citadas anteriormente e o LSM para ajuste de um modelo linear.

### 1.2.2 Organização da dissertação

Esta dissertação está dividida em cinco capítulos. No Capítulo 1, foi apresentada uma contextualização do problema de modelagem do sistema de arrefecimento de motores a combustão interna Diesel aplicados em caminhões.

O Capítulo 2 apresenta uma descrição do sistema a ser abordado pelas técnicas de modelagem de modelagem e modelo linear usando LSM.

O Capítulo 3 detalha os conceitos sobre técnicas de modelagem e técnicas de modelagem de AM empregadas nesta dissertação.

O Capítulo 4 apresenta detalhes da preparação dos dados adquiridos e os trajetos que foram utilizados, bem como ajustes específicos de cada técnica de AM durante o processo de modelagem.

No Capítulo 5 os resultados em termos de critérios de desempenho dos modelos gerados a partir de dados adquiridos no veículo de referência, nos seus ciclos de trabalho, são analisados.

Finalizando, no Capítulo 6, a conclusão da dissertação é apresentada, bem como perspectivas de futura pesquisa.

## 2 FUNDAMENTOS BÁSICOS E REVISÃO BIBLIOGRÁFICA

Neste capítulo são apresentados o motor a combustão interna como gerador de energia, o sistema de arrefecimento base para o desenvolvimento deste trabalho e na sequência um breve histórico dos trabalhos recentes e relacionados ao tema desta dissertação.

### 2.1 O MOTOR A COMBUSTÃO INTERNA

O motor a combustão interna foi idealizado para realizar trabalhos mecânicos graças à energia gerada pela combustão em seus cilindros, transferida ao virabrequim e deste para as rodas do veículo por meio da caixa de transmissão e eixos. Dessa forma o motor a combustão interna é capaz de gerar energia cinética a partir do efeito da combustão realizada dentro de seus cilindros (Heywood, 1988). Quanto maior a quantidade de combustível queimado, maior é o calor gerado, sendo assim pode-se estabelecer uma relação entre o combustível consumido e o calor gerado pela queima do mesmo. De acordo com Heywood (1988), cerca de 35% da energia química total produzida pela queima do combustível no motor é consumida como trabalho no eixo de manivela, 30% vão para os gases de escape sendo assim aproximadamente um terço dessa energia total é dissipada na forma de calor do sistema.

Segundo Iskandar e Adade Filho (2013), o desenvolvimento dos motores a combustão interna nos dias atuais é demandado essencialmente em função da redução dos poluentes na atmosfera e da redução do consumo de combustível. Estas duas demandas são impactadas diretamente pelo funcionamento adequado do motor em função do gerenciamento da temperatura de trabalho do mesmo. Esse gerenciamento é controlado eletronicamente por um *software* embarcado que controla todos os sistemas de funcionamento do motor, dentre os quais o sistema de arrefecimento, graças aos sensores, atuadores e componentes eletromecânicos que compõem um motor.

Conforme Pang et al. (2004), a temperatura de operação em motores Diesel influencia diretamente na formação de poluentes como NO<sub>x</sub> (Óxido de Nitrogênio), CO (Monóxido de Carbono), HC (Hidrocarboneto) e PM (Material

Particulado), sendo assim, um método de controle de emissões eficiente deve envolver o controle da temperatura do sistema de arrefecimento do motor, portanto o uso de um sistema eletronicamente controlado oferece a oportunidade de alterar a calibração dos parâmetros do motor para uma melhor relação entre a economia de combustível e as emissões de poluentes.

## 2.2 CIRCUITO DE ARREFECIMENTO

Os motores a combustão interna não conseguiriam funcionar sem um sistema de arrefecimento, pois, o aquecimento do motor é cíclico. Segundo Khodabakhshian et al. (2014) é importante manter o motor em temperaturas mais baixas para proteger as partes do motor e para que o fluido lubrificante mantenha suas propriedades, por outro lado, é desejável que o motor opere o mais quente possível de forma a maximizar sua eficiência térmica. Segundo BEHR (2013), esse calor aumenta à medida que o motor entra em operação, pelo aumento da demanda de torque requerida em seu eixo e por todos os acessórios dependentes do motor para funcionar. A medida que a tecnologia embarcada se torna mais complexa e as normas de emissões se tornaram mais rígidas ao longo dos últimos anos, os acessórios como direção hidráulica, ar condicionado e sistemas de controle de emissões dos gases de escape demandam ainda mais torque do motor e conseqüentemente aumenta-se a temperatura de trabalho do mesmo. Conforme mencionado em Khodabakhshian et al. (2014), aumenta-se o consumo de combustível devido à estes subsistemas auxiliares do motor .

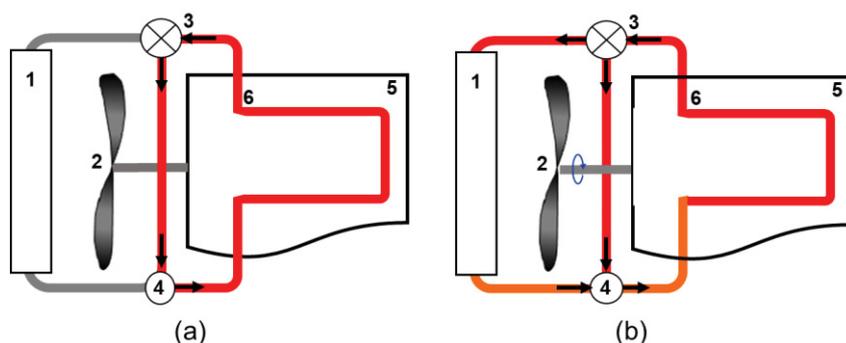
O sistema de arrefecimento é um tanto complexo e apresenta comportamento não linear. Segundo Khodabakhshian et al. (2014), apesar de ser composto de partes distintas e fáceis de se entender (conforme ilustrado na Figura 2.1), como um ventilador que ao ser ligado gera uma corrente de vento para refrigerar algo, porém quando combinados na aplicação de motor de combustão, pode ser mais complexo do que se imagina. Segundo Iskandar e Adade Filho (2013), a diminuição do tempo de ativação do ventilador em um sistema de arrefecimento controlado, resulta em maior economia de energia e em menor consumo de combustível.

Existem muitas pesquisas nesta área focando alguns dos componentes do sistema tal como a da bomba d'água, mencionado em Iskandar e Adade Filho (2013). Esta dissertação busca uma solução para auxiliar numa possível melhoria no desempenho do sistema de arrefecimento e conseqüentemente no consumo de combustível do veículo graças à modelagem da rotação ventilador do sistema e da temperatura do fluido.

Conforme ilustrado na Figura 2.1, os componentes básicos que compõem o sistema são os seguintes:

1. Radiador – troca o calor entre fluido e ar no circuito de arrefecimento.
2. Ventilador (*fan*) – responsável pelo resfriamento adicional ao do radiador.
3. Válvula termostática – controla a passagem do fluido entre o sistema e o radiador (*by-pass*).
4. Bomba d'água – responsável pela circulação do fluido por todo o sistema de arrefecimento.
5. Motor a combustão – fonte geradora de calor a ser controlada.
6. Circuito de arrefecimento – tem um papel importante na circulação do fluido no sistema.

Figura 2.1 – Diagrama do circuito de arrefecimento fechado (a), aberto (b)



FONTE: Adaptado de Bosch (2004), o autor (2018).

No momento de partida do motor, o circuito de arrefecimento permanece fechado em seu caminho original conforme apresentado na Figura 2.2(a), proporcionando um rápido aquecimento até seu ponto de funcionamento aquecido em torno de 90 graus. No entanto para manter a temperatura nesta condição a

válvula termostática circuito de arrefecimento se abre como visto na Figura 2.2(b), permitindo que o fluido circule também pelo radiador refrigerando o mesmo. Em certas condições de trabalho, com altas demandas impostas pelos sistemas do veículo, torna-se necessário uma ventilação adicional no motor realizada pelo ventilador que junto com o radiador são responsáveis por garantir que o motor não se sobreaqueça. Quando a temperatura do motor cai, a válvula termostática volta a se fechar e o fluido de arrefecimento volta a circular pelo circuito (a) da Figura 2.2, esse processo acontece durante todo o tempo em que o motor estiver rodando e sua temperatura de trabalho se altere, o sistema de arrefecimento permanece assim mantendo o motor em condições seguras de operação.

### 2.3 CONTROLE DE ACIONAMENTO DO VENTILADOR

Conforme Bosch (2004), os veículos automotores requerem capacidade substancial de arrefecimento em baixas velocidades, a ventilação por ar forçado é também necessária para o radiador. Neste cenário, os ventiladores de plástico têm essa função, visto que possuem uma capacidade de potência de até 30kW para veículos leves, contudo apresentam níveis consideráveis de ruído quando em altas rotações.

Como o motor está em constante mudança de regime de trabalho durante seu ciclo de funcionamento, o sistema de arrefecimento deve igualmente reagir no momento correto quando demandado. A reação térmica do sistema não é linear e nem instantânea, sendo que a temperatura é uma grandeza que possui atraso em suas variações, além das pressões e condições de carga a que estão sujeitas um motor. Tal panorama torna complexo o cálculo que a ECU tem de executar da demanda adequada de ventilador, de modo que a temperatura consiga se manter em condições aceitáveis para a operação do motor.

O processo de acoplar e desacoplar o ventilador não se dá de maneira instantânea causando uma histerese no controle, conforme ilustrado na Figura 2.2. A condição de repouso do ventilador é sempre em movimento em torno de 300 rpm (rotações por minuto), porém sem nenhuma conexão com a rotação do motor.

Figura 2.2 – Ciclos de acoplamento e desacoplamento do ventilador e histereses



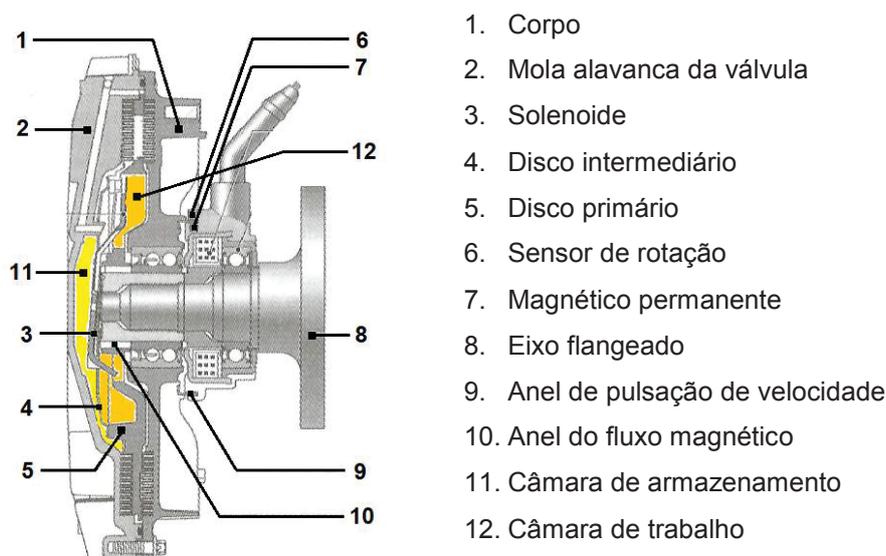
FONTE: Dados de medição em caminhão com motor Diesel, Abril/2018

Os motores pesados possuem um tipo de ventilador com embreagem acionada por fluido viscoso de silicone. O acoplamento da embreagem deste tipo de ventilador pode ser realizado por meio de um elemento bimetálico, o qual depende apenas da temperatura do próprio fluido viscoso para permitir o acoplamento da embreagem. O acoplamento pode também ser realizado por meio de um controle eletrônico com acionamento por válvula eletromagnética no lugar do bimetálico. O ventilador com controle eletrônico é o adotado para obtenção dos dados do sistema de arrefecimento, nesta dissertação.

No controle eletrônico do ventilador a ECU envia uma solicitação para que o atuador eletromagnético acione uma “embreagem”, conhecida como Visco<sup>®</sup> de forma a permitir a passagem do fluido viscoso entre as câmaras da embreagem consequentemente acionando o ventilador. Conforme apresentado na Figura 2.3, a embreagem Visco<sup>®</sup> é composta de partes móveis do lado do ventilador e do lado motor a combustão. Estas partes móveis não possuem contato direto entre si, mas são acopladas graças ao atrito do fluido viscoso que se espalha proporcionalmente por toda extensão das partes móveis da embreagem, pelo efeito da força centrífuga do ventilador entrando em movimento e fazendo com que a rotação do ventilador se iguale a rotação do motor, isto depende da quantidade, calculada pela ECU, de fluido viscoso a passar da câmara de armazenamento para a câmara de trabalho, conforme apresentado pelos itens 11 e 12 da Figura 2.3.

Em veículos comerciais, em especial os veículos com motores a diesel, o ventilador não é acionado diretamente pelo eixo do motor a combustão, mas por correias montadas junto ao eixo do motor e ao eixo flangeado do ventilador. A rotação do ventilador é em torno de 1,5 vezes a rotação do motor como pode ser observado na Figura 2.2. Uma vez acoplado o ventilador permanecerá trabalhando de acordo com as demandas de rotação impostas pela ECU até que está envie novo comando para que a válvula eletromagnética se feche e o ventilador se desacople. O desacoplamento é realizado de maneira inversa, semelhante ao processo de acoplagem, dependendo do quão rápido o fluido viscoso retorne para sua câmara de armazenamento.

Figura 2.3 – Embreagem Visco® para ventiladores automotivos



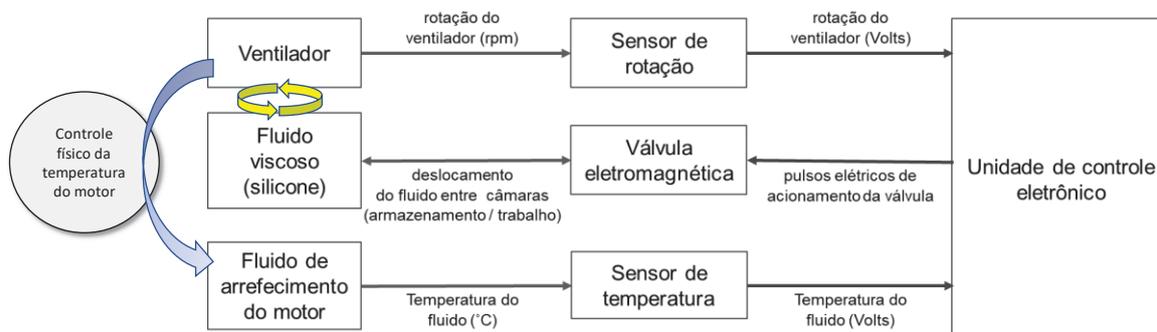
Fonte: Adaptado de Bosch (2004)

Como comentado anteriormente, o desacoplamento não acontece imediatamente, sendo assim, o ventilador permanece rodando influenciado pela rotação do motor, até que se desacople por completo. Este período é indesejado, pois, pode comprometer a eficiência do motor por interferir em seu resfriamento ou aquecimento no momento inadequado, pode também comprometer o processo de tratamento dos gases de escape gerados pelo motor Diesel, além da presença do ruído do ventilador em movimento, causando um desconforto ao motorista ao longo do tempo.

O impacto do consumo de combustível tem de ser considerado, visto que em condições normais de trabalho, esse ventilador chega a consumir em torno de 100kW, energia essa retirada do trabalho desenvolvido pelo motor a combustão. A quantidade de combustível injetada determinará a quantidade de torque entregue no eixo do virabrequim, descontando-se as perdas de energia impostas ao sistema, sendo assim o ventilador quando acoplado vai demandar do motor mais torque e conseqüentemente maior será quantidade de combustível consumido.

A ECU do motor determinará o momento de abertura ou fechamento da válvula que controla o fluxo do fluido viscoso bem como o tempo em que permanecerá acionada, conseqüentemente controlando a rotação do ventilador. O retorno do controle baseia-se em um sensor que medirá a rotação do ventilador (detalhes na Figura 2.4). A temperatura do motor também é utilizada no cálculo da rotação do ventilador e geralmente é representada pela a temperatura do fluido de arrefecimento ou ainda pela temperatura do óleo lubrificante do motor.

Figura 2.4 – Diagrama do ventilador com controle eletrônico



FONTE: Adaptado de Bosch (2004), o autor (2018)

## 2.4 REVISÃO BIBLIOGRÁFICA

Neste capítulo são apresentados trabalhos realizados cujos objetivos foram a modelagem do sistema de arrefecimento e seus diversos componentes e diferentes abordagens ao longo dos anos, desde as mais antigas quanto as mais recentes.

Rio (1976) registrou a patente de um sistema de acionamento do ventilador, uma vez que a energia demandada deste acionamento para a refrigeração de motores a combustão reduz a potência fornecida pelo motor e assim reduzindo drasticamente seu desempenho. Consequentemente deseja-se que a rotação do ventilador seja controlada pela demanda de resfriamento do motor independente da rotação do motor a combustão e velocidade do veículo. O conceito refere-se a um mecanismo de acionamento e controle do sistema de arrefecimento pelo controle da temperatura e rotação do ventilador, sendo composto de uma bomba hidrostática acionada pelo próprio motor para transmitir energia necessária ao acionamento do ventilador, o mecanismo entrega um acionamento mecânico como saída controle.

Pang et al. (2004) analisaram a influência do controle da temperatura do fluido de arrefecimento de motores Diesel no controle dos gases emitidos por este tipo de motor. Os autores realizaram testes de motor com ciclos de emissões específicos. Neste contexto, um programa computacional foi gerado para validar a proposta, que permitiu demonstrar as mudanças nas condições térmicas do controle de temperatura do fluido de arrefecimento na prática. Os fatores considerados no programa de teste foram a temperatura do fluido de arrefecimento, o tempo de injeção e a taxa da válvula de recirculação dos gases de escape (EGR).

Campos (2010) apresentou a aplicação de nano fluidos no sistema de arrefecimento de motores a combustão interna. “*Os nano fluidos são suspensões de partículas menores que 100nm de fluidos convencionais de troca de calor, tais como água, óleo, etileno glicol, entre outros. Devido a suas características favoráveis de transferência de calor.*” O trabalho se desenvolveu através da elaboração de modelos matemáticos para a operação do motor em regimes permanente e transiente para avaliar seu processo de aquecimento. Foi utilizado o *software* de engenharia EES (*Engineering Equation Solver*) para simulações

considerando-se os seguintes componentes do sistema de arrefecimento automotivo: radiador, camisas do bloco de cilindros, termostato e bomba do fluido de arrefecimento.

Iskandar e Adade Filho (2013) realizaram pesquisa do controle de arrefecimento de um motor Diesel, visando a redução de emissões e do consumo de combustível pela análise e projeto de um sistema substituto ao tradicional sistema de arrefecimento (*cooling*) do motor Diesel, o sistema em estudo é composto por uma bomba de água elétrica, um ventilador acionado eletricamente, uma válvula termostática controlada eletronicamente e um algoritmo de controle. Este conjunto visa propiciar um controle aprimorado da temperatura de água do motor Diesel com objetivo na redução do consumo de combustível. Os ajustes do controlador foram calculados por meio de algoritmos de otimização e co-simulação computacional utilizando o *software* GT-Tool (marca registrada da *Gamma Technologies*) e ambiente computacional Matlab/*Simulink*<sup>®</sup> (marca registrada da *Mathworks*). Resultados e testes realizados na abrangência deste trabalho mostraram ser possível prever que o uso de um sistema de arrefecimento em motores eletronicamente controlados tende à redução do consumo de energia além da redução da temperatura nos componentes a óleo do motor.

Sandu (2016) investigou o aumento de desempenho de motores Diesel carga pesada quando o ventilador, com acoplamento viscoso, do sistema de arrefecimento é acionado descontinuamente ao invés de estar acoplado diretamente ao eixo do motor. O teste foi realizado em duas etapas, primeiro num banco de testes de fluxo de ar por perda de pressão e depois diretamente montado em um motor Diesel turbo carga pesada. O objetivo foi fazer com que o ventilador fosse acoplado apenas quando realmente necessário, uma vez que o ventilador viscoso tem rotação independente a do eixo do motor, sendo assim controlado diretamente pelas características de arrefecimento do motor. O estudo demonstrou que se pode ter um ganho de potência próximo ao descrito na literatura, na ordem de 5% de economia de combustível.

Romanov e Khozeniuk (2016) investigaram o sistema de arrefecimento do motor utilizando modelos preparados com a ferramenta computacional de dinâmica dos fluidos (*Computational Fluid Dynamics – CFD*), observando os diversos modos de início de circulação do fluido pelos seus canais dentro do bloco do motor, de

modo a se obter estimativas quantitativas das superfícies de resfriamento de vários cilindros.

Zhiqiang et al. (2016) apresentaram o problema do alto consumo de potência e baixa taxa de utilização de energia de ventiladores com acionamento hidráulico em veículos militares. Por meio de um sistema *fuzzy* para os controladores PID (Proporcional, Integral e Derivativo) do ventilador procurou-se uma melhora para o sistema. Através dos experimentos pode-se estabelecer a influência do torque e rotação do motor na rotação do ventilador e assim conseguir uma melhor adaptação destes ao ventilador proporcionando um controle eficiente da rotação do mesmo.

Haghighat et al. (2018) apresentaram um modelo de controle do sistema de arrefecimento de modo a reduzir o consumo de combustível e emissão de poluentes. O modelo contou com a utilização de componentes controlados do sistema de arrefecimento, como bomba de água elétrica, ventilador elétrico e termostato aquecido. Este trabalho mostrou que o controle da temperatura do motor impactou numa redução de 1,1% no consumo de combustível para um motor de 1,4 litros em todas as partes do ciclo de emissões NEDC (*New European Driving Cycle*), incluindo a parte de aquecimento do motor bem como condições de baixas e altas cargas. Dentre alguns estudos apresentados com estratégias alternativas para resfriamento do motor obteve-se redução do consumo de combustível em torno de 2 a 3% até 5% dependendo da estratégia adotada.

### 3 FUNDAMENTOS DE MODELAGEM

Neste capítulo são abordados os fundamentos de modelos matemáticos úteis para a modelagem do sistema de arrefecimento além das técnicas de regressão da área de AM denominadas ANN e SVR, além de modelo linear com ajuste de parâmetros via mínimos quadrados em batelada.

#### 3.1 MODELAGEM DE UM SISTEMA

Um sistema é um conjunto de várias partes ou elementos ou ainda funções que juntas realizam uma finalidade, um propósito específico, que não é possível quando separadas. De acordo com INCOSE (2011), estes elementos podem incluir pessoas, *hardware*, *software*, os meios, políticas e documentos necessários para se produzir resultados de um sistema. Estes resultados incluem diferentes níveis sistêmicos de qualidade, propriedades, características, funções, comportamentos e desempenhos. O relacionamento entre estas partes e como são interconectadas compõem um sistema. Segundo Ljung e Söderström (1983), o conhecimento das propriedades de um sistema é chamado de modelo, o qual pode ser dado por diferentes formas, como por exemplo modelos mentais ou intuitivos, modelos gráficos e modelos matemáticos.

Segundo Aguirre (2015) a Identificação de sistemas trata de um sistema alternativo no qual propõe-se obter um modelo matemático que explique em parte e de forma aproximada a relação de causa e efeito presente nos dados. O modelo matemático de um determinado sistema pode então ser construído a partir da observação de dados fornecidos deste sistema. Segundo Ljung (1999), um sistema é um objeto no qual variáveis de diversos tipos interagem entre si e produzem sinais observados, estes são as saídas de interesse no processo. Quando se assume uma certa relação entre as variáveis observadas se estabelece um modelo do sistema. Um modelo nunca será a verdade final de um sistema, mas pode ser descrito como sendo “bom o bastante” sob certos aspectos do sistema e de interesse específico”. Uma vez treinado o algoritmo estará apto a prever valores na saída y para qualquer entrada x do sistema que foi “identificado”.

Neste processo parte dos dados de entrada será utilizada para o treinamento do modelo e parte será empregada na sequência para a validação do

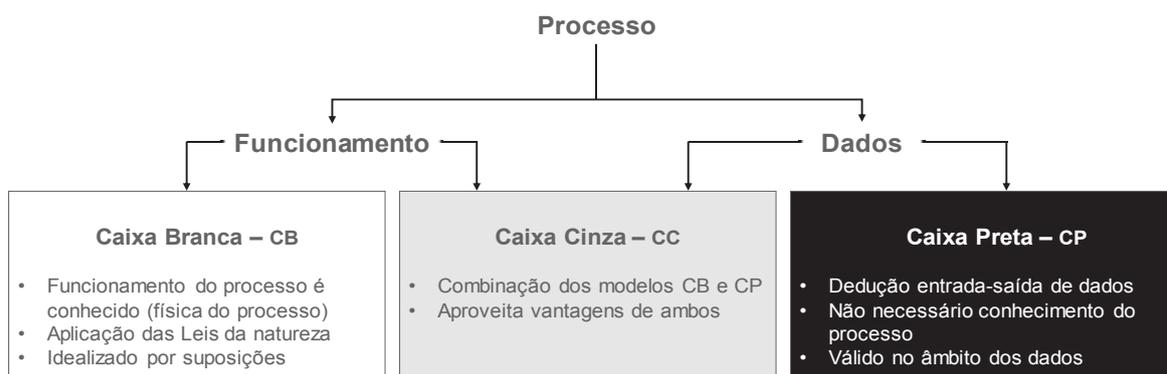
mesmo. A validação mostra quanto um modelo é capaz de representar os dados observados de um sistema, se ele é adequado o bastante para o propósito daquele sistema inicialmente treinado. Segundo Lima (2004), a capacidade de aproximação, geralmente vinculada a modelos matemáticos, é uma propriedade universal e sua aplicação a problemas práticos requer um grau de flexibilidade do modelo tanto para treinamento supervisionado como para o não-supervisionado a partir de dados amostrados. Convencionou-se chamar de maximização da capacidade de generalização, ou seja, o desempenho ótimo junto a dados não observados durante o processo de treinamento. A SVM é uma das técnicas de AM cujo embasamento teórico possui aplicação na maximização da capacidade de generalização de modelos de previsão.

### 3.1.1 Classificação de modelos

Segundo Zendeboudi et al. (2018), a classificação de um modelo depende da filosofia e arquitetura do modelo no qual variáveis de entrada e saída estão relacionadas, sendo assim baseia-se na extensão da confiabilidade do modelo pelo funcionamento ou conhecimento do processo, no relacionamento dos dados de entrada e saída ou na combinação de ambos. Dentre as várias aplicações dos modelos estão, simulação, otimização, estimação, controle e monitoramento.

Segundo Nelles (2001), três são as abordagens básicas de modelos na teoria de modelagem, conforme Zendeboudi et al. (2018), baseiam-se na classificação fundamental em termos de dependência do conhecimento do processo versus os dados do processo. As três classes de modelos são caixa branca, caixa cinza e caixa preta conforme Figura 3.1.

Figura 3.1 – Classes de modelos e sua confiabilidade, funcionamento x dados



Adaptado de Zendehboudi et al. (2018)

- Modelos caixa branca (*white box*): baseados em leis e princípios básicos da Física, Química entre outros, onde as equações e parâmetros podem ser determinados teoricamente, não dependendo de dados, embora possa usá-los em algum momento para validar o modelo.
  - Desvantagem: o modelo pode ser de difícil obtenção.
- Modelos caixa preta (*black box*): baseiam-se exclusivamente nos dados fornecidos, sendo modelo e parâmetros determinados a partir de dados experimentais de maneira empírica, não se tem conhecimento prévio sobre o sistema a ser modelado e não dependem dos princípios básicos teóricos. Conforme Aguirre e Corrêa (2004), a escolha da representação e da sua estrutura é feita de forma empírica nos casos mais simples. Em casos mais complexos, como por exemplo na identificação de sistemas não-lineares, esta escolha é crítica o que justifica o uso de métodos mais sofisticados para seleção da estrutura do modelo. Tais métodos baseiam-se na maior parte das vezes em técnicas de Álgebra linear e em conceitos de Estatística.
  - Desvantagem: o número excessivo de parâmetros gerados.
- Modelos caixa cinza (*grey box*): podem ser considerados como sendo uma combinação entre os modelos caixa branca e os modelos caixa preta, sendo assim, o conhecimento dos princípios básicos ou informações prévias sobre os dados medidos podem ser utilizados. Conforme Aguirre e Corrêa (2004), a identificação caixa-cinza é uma área bastante ampla e por se tratar de um assunto relativamente novo muitos

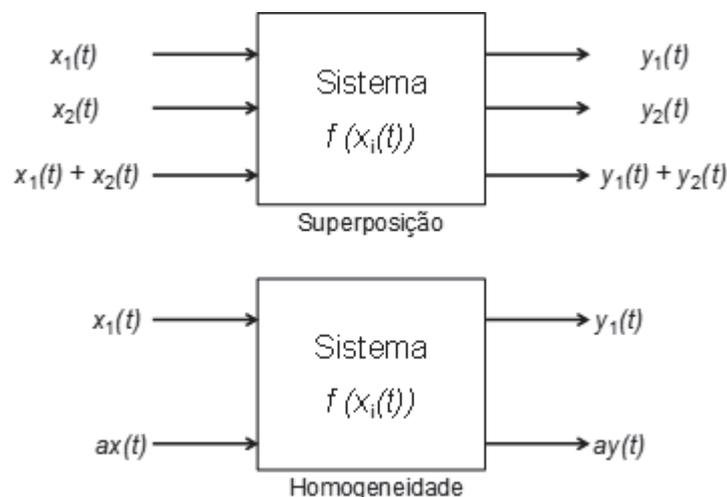
problemas estão praticamente em aberto. Segundo Ljung (1999), os modelos com parâmetros ajustáveis com interpretação das leis da Física são chamados caixas cinza.

### 3.1.2 Sistemas lineares e não lineares

Os modelos são caracterizados como sendo lineares ou não-lineares, possuem propriedades específicas e algumas comuns entre ambos.

Os sistemas lineares são caracterizados por suas saídas dependerem linearmente de suas entradas seja em tempo,  $t$ , contínuo ou discreto. Essa linearidade baseia-se no princípio da superposição. Um sistema linear que atende ao princípio de superposição baseia-se em duas propriedades: aditividade e homogeneidade, sendo aditividade onde cada saída deve corresponder a uma entrada específica, por exemplo  $y_1(t)$  corresponde a entrada  $x_1(t)$ ,  $y_2(t)$  à  $x_2(t)$  e assim sucessivamente. O sistema estará sujeito a aditividade se  $x(t) = (x_1(t) + x_2(t))$  então  $y(t) = (y_1(t) + y_2(t))$ . Na propriedade de homogeneidade, cada saída  $y(t)$  deve corresponder a uma mesma escala ou fator de escala “a” quando aplicado a sua respectiva entrada  $x(t)$ , assim  $ax(t)$  corresponderá à saída do modelo  $ay(t)$ , conforme apresentado na Figura 3.2.

Figura 3.2 – Modelo de sistemas lineares: superposição, aditividade e homogeneidade



FONTE: Adaptado de Ogunfunmi (2007)

Geralmente sistemas lineares são invariáveis no tempo e são conhecidos como LTI (do inglês *Linear Time-Invariant*), cujas características ou propriedades não variam ou mudam com o tempo, dessa forma um sistema é dito invariável no tempo quando o deslocamento de suas entradas  $x(t-t_0)$  no tempo se repete igualmente em sua saída  $y(t-t_0)$ , dito princípio de causa-efeito. Segundo Aguirre (2015) a invariância temporal implica que o comportamento do sistema modelado não varia com o tempo, isso não significa que as variáveis do sistema têm seus valores constantes. Na verdade, os valores das variáveis do sistema flutuam com o tempo. Essa evolução no tempo é normalmente regida por uma lei a qual determina a dinâmica do sistema.

Outras propriedades de sistemas lineares são a casualidade, estabilidade e memória. Conforme Ogunfunmi (2007), um sistema é causal se sua saída depende apenas das entradas presentes ou passadas mas não das futuras. Sistemas não causais podem existir em função de atrasos (memória) mas não em tempo real. Um

sistema é estável se apresentar uma entrada  $x(t)$  limitada (*bounded-input*) que induzirá a uma saída  $y(t)$  também limitada (*bounded-output*). Este tipo de sistema é conhecido como BIBO (*bounded-input bounded-output*). Um sistema sem memória é aquele cuja saída em determinado instante de tempo depende exclusivamente de sua entrada naquele mesmo instante, se a saída no tempo inicial  $t_0$  depender de entradas ocorridas antes ou depois do tempo inicial  $t < t_0$  ou  $t > t_0$  o sistema é dito com memória.

Os sistemas não-lineares são os que não obedecem de alguma forma aos princípios de linearidade, ou seja, as saídas são funções não-lineares de suas entradas. Segundo Ogunfunmi (2007), muitos sistemas polinomiais obedecem ao princípio da superposição mas não ao de homogeneidade, portanto são não-lineares. Algumas vezes estes sistemas possuem propriedades lineares, mas são variantes no tempo, dessa forma apresentam um comportamento não linear o qual leva a sistemas não causais. É também verdadeiro que para sistemas lineares LTI as componentes de frequência presentes em um sinal de saída são as mesmas presentes nas entradas de um sistema. No entanto, em sistemas não-lineares as frequências da saída não são exatamente as mesmas presentes na entrada, pois, existem sinais de outras frequências. Quando esses sinais são discretizados no tempo as saídas de um sistema não-linear podem estar sujeitas à componentes

*aliased* (sobrepostos) da entrada os quais necessitam ser amostrados duas vezes a taxa da maior frequência do sinal de entrada, como definido por Nyquist, para se evitar tal efeito.

Existem vários tipos de não-linearidade que podem ser modeladas por meio de polinômios que permitem estender comportamentos de sistemas lineares para os sistemas não-lineares. Representar os sistemas não-lineares implica em generalização das propriedades enumeradas para sistemas lineares cuja relação entrada-saída poderá implicar em causalidade nos sistemas homogêneos de grau 'n', conforme citado em Ogunfunmi (2007), sendo que uma soma infinita de termos homogêneos é conhecida por séries polinomiais de Volterra que permitem representar a não linearidade por meio de polinômios.

### 3.1.3 Tipos de aprendizado

Lorena, Ana C. e Carvalho (2003) citam o aprendizado de máquina (AM) como um campo de pesquisa da inteligência computacional que estuda o desenvolvimento de métodos capazes de extrair conhecimento a partir de dados amostrados por meio de algoritmos que geram classificadores para os dados apresentados, assim podem prever classes de instâncias quaisquer do domínio em que foi treinado. Segundo Lorena e Carvalho (2007), na modelagem o objetivo é desenvolver algoritmos capazes de fazer uma estimativa com o menor erro possível, diz-se para aprendizado de máquina que estes algoritmos possuem habilidade de generalização sendo que:

Generalização é a capacidade de prever corretamente uma classe de novos dados tornando-os possíveis de serem separados. (Fonte: Lorena e Carvalho, 2007).

Generalização no contexto de algoritmos de aprendizado significa um problema em que os dados de entrada com características ou padrões semelhantes sejam corretamente classificados.

O Aprendizado de máquina pode ser dividido em três grandes grupos de técnicas e problemas que cada um consegue resolver, sendo eles, aprendizado supervisionado, semi-supervisionado e não-supervisionado suas principais características são mencionadas no Quadro 3.1.

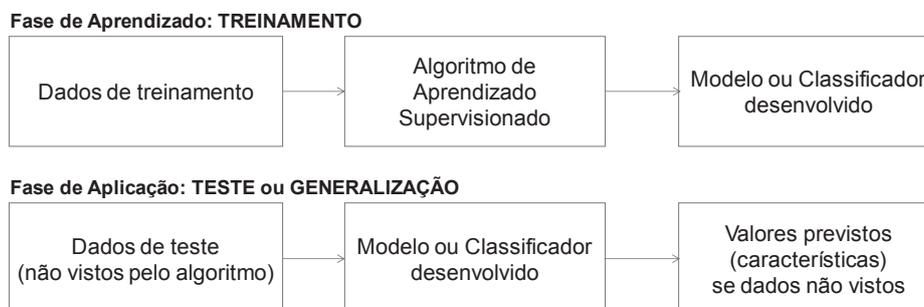
Quadro 3.1 – Técnicas de aprendizado de máquina

Técnica	Algoritmos	Tarefa de aprendizado	Exemplos de áreas de aplicação
<b>SUPERVISIONADA</b> O termo supervisionado vem do fato de as características dos dados de treinamento atuarem como professores os quais ensinam os algoritmos.	<u>Classificação:</u> - Reconhecimento de padrões	Construir classificadores que possam classificar dados não vistos ainda.	Reconhecimento de objetos e de escrita manual, análises de doenças.
	<u>Regressão:</u> - Funções de aproximação	Encontrar um mapeamento entre entrada e saída, ou seja, uma função que represente $N$ vetores de entrada $x_i \in \mathbb{R}^n$ e a saída sendo um escalar $x_i \in \mathbb{R}$ .	Análise de séries temporais, sistemas de controle, de navegação e previsão na área financeira.
<b>SEMI-SUPERVISIONADA</b> Dados existentes, $x_i \in \mathbb{R}^n$ , porém com algumas de suas características $y_i$ .	<u>Classificação:</u> - Análise de componentes principais ( <i>Principal Component Analysis</i> , PCA) ou independentes ( <i>Independent Component Analysis</i> , ICA) - Fatoração de matrizes não-negativas - Agrupamento de dados ( <i>clustering</i> )	Prever características de dados não categorizados pelo uso total dos dados fornecidos, tende a ser caro. Treinamento de dados consiste em ambos dados, categorizados e não-categorizados.	Processamento de ditados, categorização de textos e bioinformática.
<b>NÃO-SUPERVISIONADA</b> Apenas os dados estão disponíveis, $x_i \in \mathbb{R}^n$ sem suas características $y_i$ correspondentes, (sem professor em forma de características).	<u>Classificação, Regressão:</u> - ICA, PCA entre outros - Agrupamento de dados ( <i>clustering</i> )	Conceito denso e que pode ser visto de diferentes perspectivas. Como resultado tem-se uma "nova" representação ou explanação dos dados observados.	PCA em medições cerebrais, eletroencefalografia, sistemas de comunicação <i>wi-fi</i> com múltiplas antenas.

Fonte: Huang et al. (2010)

De maneira simples todas as três técnicas dependem dos dados fornecidos por um determinado sistema para aprender com os chamados dados de treinamento, conforme ilustrado na Figura 3.3.

Figura 3.3 – Fases de aprendizado supervisionado



Fonte: Adaptado de Huang et al. (2010)

### 3.2 MÉTODO DE REGRESSÃO LINEAR

Segundo Ljung (1999), um modelo é dito de regressão quando consegue prever resultados baseados em amostras de variáveis ocorridas no tempo passado, a quais representam um sistema discretizado no tempo em que seus dados são sempre coletados por amostragem, regida por um período de amostragem, representado pela equação das diferenças linear:

$$y(t) = -a_1y(t-1) - \dots - a_ny(t-n) + b_1u(t-1) + \dots + b_mu(t-m) \quad (3.1)$$

onde  $y$  é a saída do sistema,  $u$  é entrada do sistema,  $a$  e  $b$  são os polos e zeros discretos, e  $n$  e  $m$  são as ordens, ou seja, o número de polos e zeros dos modelos, respectivamente.

Representando-se por vetores tem-se os seguintes valores:

- O vetor de parâmetros dado por

$$\theta = [a_1 \dots a_n \ b_1 \dots b_m]^T \quad (3.2)$$

- O vetor de sinais observados ou vetor de dados de um sistema onde

$$\varphi(t) = [-y(t-1) \dots -y(t-n) \ u(t-1) \dots u(t-m)]^T \quad (3.3)$$

Sendo assim, pode-se reescrever a equação (3.1), conhecida como modelo auto regressivo com entradas exógenas (*Auto regressive with exogenous inputs – ARX*), sendo um modelo linear de regressão em que  $y(t)$  é calculado baseado nas amostras passadas de (3.3):

$$\hat{y}(t|\theta) = \theta^T \cdot \varphi(t) = \varphi^T(t) \cdot \theta \quad (3.4)$$

onde  $\hat{y}$  é o vetor de saídas previstas,  $\varphi(t)$  é o vetor de dados, ou vetor de regressão e  $\theta$  é o vetor de parâmetros a ser determinado. O previsor é dado por meio do valor estimado  $\hat{y}$  correspondente ao produto escalar entre o vetor de dados  $\varphi(t)$  e o vetor de parâmetros  $\theta$ .

### 3.2.1 Método dos mínimos quadrados

O método dos mínimos quadrados (do inglês *Least Square Method* – LSM) foi primeiramente desenvolvido por Gauss in 1795 e tornou-se um dos mais aplicados na solução de problemas de otimização. No aprendizado supervisionado o erro  $e(i)$  costuma ser a diferença entre o processo medido  $y_i$  e a saída do modelo  $\hat{y}_i$  para um certo número de amostras de treinamento ( $i = 1, \dots, N$ ).

O objetivo do método é encontrar a saída do modelo que melhor se aproxime da saída do processo, isto quer dizer, com o menor erro da soma dos quadrados da função de perda. Segundo Nelles (2001), significa determinar a melhor combinação linear dos regressores pelos parâmetros de otimização  $\theta_1, \dots, \theta_n$ , assim a expressão (3.5) representa um problema de regressão linear onde em estatística os  $x_i$  são chamados regressores ou variáveis independentes e  $y$  de variável dependente,  $\theta_i$  são os chamados coeficientes de regressão. A saída prevista  $\hat{y}$  do modelo depende linearmente dos  $n$  parâmetros  $\theta_n$ , tal que.

$$\hat{y} = \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n = \sum_{i=1}^n \theta_i x_i \quad (3.5)$$

O LSM descrito por Ljung (1999) leva em consideração o erro na estimativa do modelo de regressão e o erro nos dados fornecidos, aqui mencionado como ruído branco  $e(t)$  sendo uma sequência de dados aleatórios independentes com média zero e uma determinada variância  $\sigma^2$ , assim, busca-se a qualidade da estimação pela natureza do ruído, da estrutura do modelo e dos valores medidos mais próximos possível do sistema a ser modelado. De maneira similar a (3.4) tem-se a equação a diferenças dada por:

$$\hat{y}(t) = \varphi^T(t) \cdot \theta + e(t) \quad (3.6)$$

Supondo-se um sistema cujos valores dos parâmetros em  $\theta$  são desconhecidos, porém com seus dados de entrada  $u(t)$  e saída  $y(t)$  armazenados no vetor  $Z^N$  que os representam num intervalo de tempo  $1 \leq t \leq N$ :

$$Z^N = \{u(1), y(1), \dots, u(N), y(N)\} \quad (3.7)$$

O LSM propõe que se busque um valor que minimize a equação:

$$V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N [y(t) - \varphi^T(t)\theta]^2 \quad (3.8)$$

O processo de minimização se dá encontrando um valor  $\theta$  que minimize  $V_N$  por  $\hat{\theta}_N$ :

$$\hat{\theta}_N = \arg \min_{\theta} V_N(\theta, Z^N) \quad (3.9)$$

Uma vez que  $V_N$  é quadrático em  $\theta$ , pode-se encontrar o valor mínimo de  $\theta$  para minimizar  $V_N$  pela derivada de  $\theta$  para zero:

$$0 = \frac{d}{d\theta} V_N(\theta, Z^N) = \frac{2}{N} \sum_{t=1}^N \varphi(t)(y(t) - \varphi^T(t)\theta) \quad (3.10)$$

Assim tem-se a equação dos mínimos quadrados, tal que:

$$\hat{\theta}_N = \left[ \sum_{t=1}^N \varphi(t)\varphi^T(t) \right]^{-1} \sum_{t=1}^N \varphi(t)y(t) \quad (3.11)$$

Ainda segundo Ljung (1999), tanto a forma simples de modelagem (3.1) quanto a dos mínimos quadrados (3.11) representam a base das equações de identificação de sistemas por métodos de identificação paramétricos e podem ser estendidos para várias entradas pela redefinição de (3.3), baseados apenas de dados medidos independentes das leis física do sistema.

Segundo Haykin (2001), o LSM além de sua simplicidade, é independente do modelo e conseqüentemente robusto, assim pequenas incertezas do modelo e pequenas perturbações resultam apenas em pequenos erros de estimativa, possui a habilidade de funcionar satisfatoriamente tanto em um ambiente estacionário como em um ambiente não estacionário em que as estatísticas variam como tempo.

As limitações principais do LSM são a taxa de convergência lenta e a sensibilidade às variações na auto-estrututa da entrada e requer um número de iterações igual a cerca de dez vezes a dimensionalidade do espaço de entrada para alcançar a condição de estabilidade, sendo crítico para altas dimensionalidades do espaço de entrada.

### 3.2.2 Mínimos quadrados recursivo

O algoritmo (3.5) aplicado na estimação de parâmetros dos mínimos quadrados recursivos é uma forma algébrica de processar as medidas dos mínimos quadrados em batelada (*batch*), cujas propriedades são:

- Polarização (do inglês *bias*): O estimador é não polarizado se os parâmetros convergem para os valores verdadeiros quando o número de iterações aumenta, sendo assim o ruído branco ou perturbação possui média nula e variância igual a um, além de saída-entrada serem estatisticamente independentes do erro. Portanto, se o estimador apesar de muitas observações a estimativa não se aproxima do valor verdadeiro é tido como polarizado.
- Covariância (do inglês *covariance*): Esta propriedade é importante para se determinar se existe solução dos mínimos quadrados, sendo assim deve ser sempre definida positiva, invertível, não singular. Conforme Nelles (2001), a matriz de covariância é dada pela equação:

$$\text{cov}\{\hat{\theta}\} = \sigma^2(X^T \cdot X)^{-1} \quad (3.12)$$

onde  $X$  é a matriz de regressão e  $X^T$  é chamada de pseudo inversa da matriz  $X$ , ambas entradas para a matriz de covariância dos parâmetros. A matriz de covariância é proporcional a  $1/N$ . Ainda segundo Nelles (2001), se as entradas  $X^T$  e  $X$  aumentam linearmente com o número de dados de treinamento  $N$ , então ao coletar-se dados suficientes qualquer ruído poderá ser compensado.

### 3.3 MÁQUINA DE VETOR DE SUPORTE

Segundo Haykin (2001), a máquina de vetor de suporte (do inglês *Support Vector Machine* – SVM) é um método de aprendizado de máquina e segue o princípio da minimização estrutural do risco dimensional fundamentado na teoria Vapnik-Chervonenkis (VC).

O método SVM é uma forma de generalização não-linear do algoritmo desenvolvido na Rússia nos anos 1960 por Vapnik-Lerner (1963), é baseado na teoria do aprendizado estatístico Vapnik-Chervonenkis (1964), tem se desenvolvido por mais de três décadas por Vladimir Vapnik e seus colaboradores Smola e Schölkopf (2004). Segundo Lima (2004), por ser fundamentado pela teoria do aprendizado estatístico, a SVM possibilita e investigação de formas de estimação de dependências funcionais com base em conjuntos de dados e permite analisar formalmente o comportamento e as limitações inerentes a algoritmos de Aprendizado de máquina. Conforme Vapnik et al. (1999), a SVM é um modelo universal de aprendizado de máquina no qual mantem-se o valor do risco empírico e minimiza-se o intervalos de confiança.

#### 3.3.1 Dimensão VC

De acordo com Haykin (2001), a construção do hiperplano de separação ótimo depende do princípio da minimização estrutural do risco que é fundamental na teoria da dimensão Vapnik-Chervonenkis (1971), ou simplesmente dimensão VC, dependendo do cálculo de um produto escalar do *kernel*. Para Lima (2004), a dimensão VC é um índice escalar que mede a complexidade intrínseca de uma classe de funções.

A teoria da convergência uniforme do risco empírico  $R_{emp}(w)$  para o risco real  $R(w)$  inclui os limites da taxa de convergência baseados no parâmetro VC, a qual representa o número máximo de exemplos de treinamento que podem ser aprendidos pela máquina sem erro, para todas as categorizações possíveis das funções de classificação, sendo assim VC é uma medida da capacidade da família de funções de classificação realizadas pelo aprendizado de máquina

A convergência em probabilidade é dada pela expressão:

$$P|a_N - a_0| > \delta \xrightarrow{P} 0 \text{ quando } N \rightarrow 0 \quad (3.13)$$

onde  $a_N$  representa uma sequência de variáveis aleatórias que convergem em probabilidade para uma variável  $a_0$  se para qualquer  $\delta > 0$  for válida a relação probabilística dada por (3.13)

Considerando-se uma regra de decisão simples como um exemplo em um espaço  $m$ -dimensional  $\mathcal{H}$  de vetores de entrada descrito como

$$f: y = \varphi(w^T x + b) \quad (3.14)$$

onde  $x$  neste caso é um vetor de pesos  $m$ -dimensional,  $b$  é a polarização (*bias*<sup>2</sup>). A função  $\varphi$  é uma função de limiar  $\varphi(v)$  vale 1 se  $v \geq 0$  e vale 0 se  $v < 0$ . A dimensão VC para a regra de decisão (3.14) é dada por

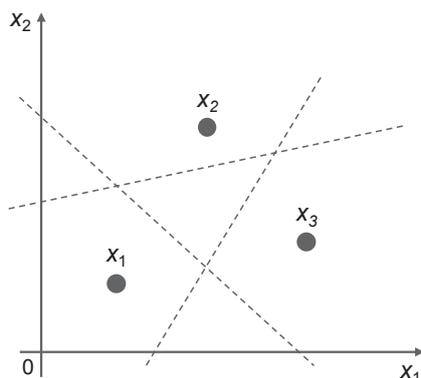
$$VC_{dim}(f) = m + 1 \quad (3.15)$$

Portanto ao obter-se um espaço de entrada bidimensional  $m=2$ , por (3.15) a dimensão será VC igual a 3, conforme ilustrado na Figura 3.4. Na Figura 3.4 existem três pontos  $x_1$ ,  $x_2$  e  $x_3$  com três possibilidades diferentes de categorização onde observa-se um número máximo de três linhas para separar os pontos.

---

<sup>2</sup> *Bias* em SVM faz parte da representação geométrica da posição de um hiperplano como apresentado na equação (3.16), diferente do bias definido como a propriedade de polarização do estimador LSM.

Figura 3.4 – Par de distribuições de dados bidimensional (dimensão VC)

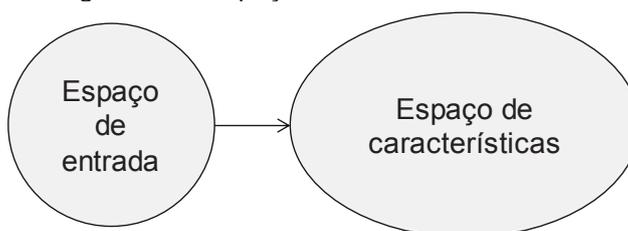


Fonte: Adaptado de Haykin (2001)

### 3.3.2 Espaço de características

A SVM busca construir um hiperplano de separação definido com uma função linear de vetores retirados do espaço de características em vez do espaço original. Como nem sempre as classes envolvidas são possíveis de se separar de maneira linear, a SVM muda os dados de entrada para um espaço de alta dimensão no qual se torna possível a separação das classes, sendo esta uma das principais características da SVM.

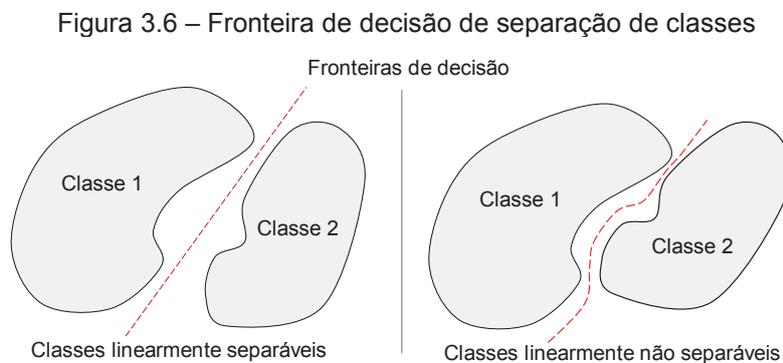
Figura 3.5 – Espaços de entrada e de características



Fonte: O autor 2018

Originalmente desenvolvida para classificação de padrões a SVM baseia-se em encontrar as superfícies de decisão de cada classe (Figura 3.6) e a partir desta etapa calcular a melhor condição de separação por meio de vetores de suporte. Conforme Haykin (2001), A SVM depende da extração de um subconjunto de dados de treinamento que serve como vetores de suporte. Os vetores de suporte são

aqueles pontos de dados que se encontram mais próximos da superfície de decisão e, portanto, mais difíceis de classificar.

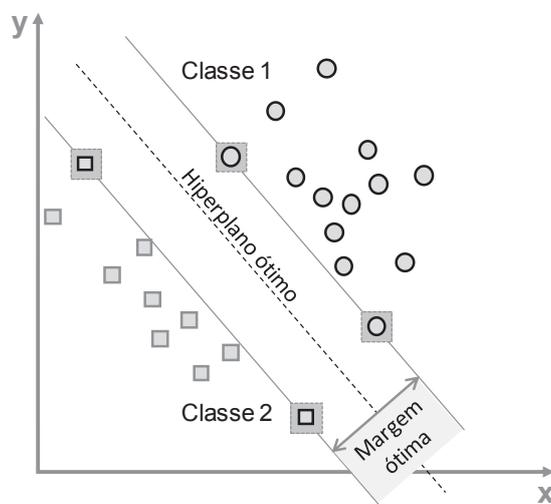


Fonte: Adaptado de Haykin (2001)

### 3.3.3 A margem de separação

Na busca da separação entre classes, segundo Cortes e Vapnik (1995), tenta-se a maximização do espaço de entrada entre os dados separáveis fornecidos num ambiente normal de duas dimensões de modo que se obtenha uma distância entre as classes com apropriada tolerância ao erro. Essa distância é conhecida em SVM como margem de separação, sendo esta ilustrada na Figura 3.7.

Figura 3.7 – Separação ótima entre duas classes (SV em cinza)



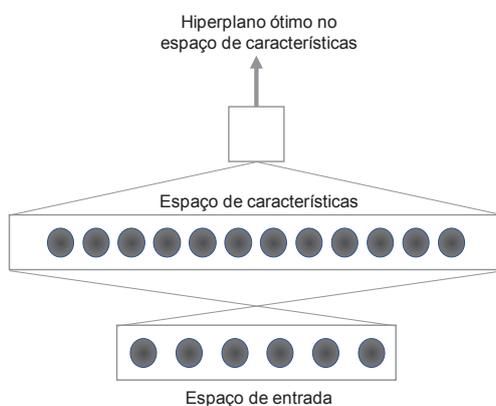
FONTE: Adaptado de Vapnik et al. (1999)

### 3.3.4 Hiperplano ótimo

Na SVM o processo de aprendizagem se dá por meio de dados apresentados como espaço de entrada mapeados em um espaço de maior dimensionalidade conhecido como espaço de características  $\mathcal{H}$ , de onde então constrói-se o hiperplano ótimo, podendo assim lidar com casos linearmente não separáveis.

A SVM busca pelo hiperplano ótimo, pelo mapeamento do espaço entrada em um espaço de características, conforme apresentado na Figura 3.8.

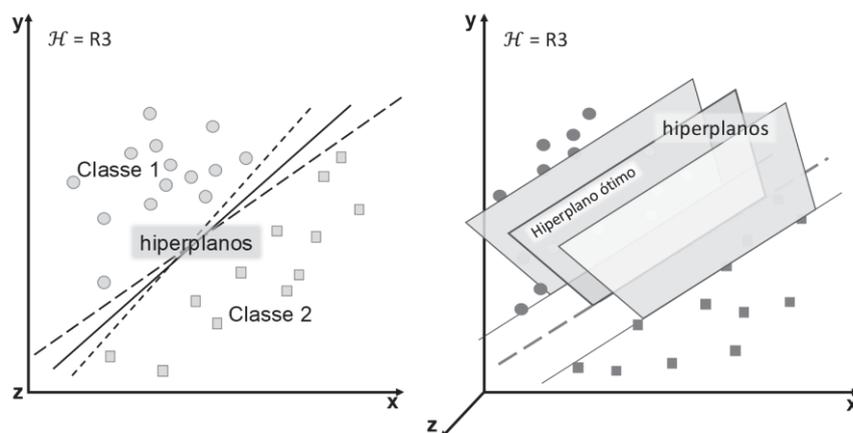
Figura 3.8 – O mapeamento de espaços na SVM



Fonte: Adaptado de Vapnik et al. (1999)

Uma das áreas de utilização da SVM é a de processamento de imagem por meio de funções que dividem os dados fornecidos ou espaços de características em regiões para o reconhecimento de padrões de acordo com o número de classes envolvidas no problema. Segundo Santos (2002), uma forma simples de particionar um espaço euclidiano de  $n$  dimensões é pela utilização de hiperplanos. De acordo com Vapnik et al. (1999), o conceito de hiperplano ótimo de separação de classes foi definido por Vapnik e Chervonenkis (1974), inicialmente para dados linearmente separáveis, sendo o conceito definido por uma função de decisão linear que buscará entre todos os possíveis hiperplanos (Figura 3.9-esquerda), qual destes possui a maior distância entre duas classes separáveis, calculando-se uma margem máxima entre as duas. A distância média na margem de separação entre as classes é denominada de hiperplano ótimo com a direita na Figura 3.9

Figura 3.9 – Hiperplanos de alta dimensão para separação de classes



Fonte: Adaptado de Vapnik, O autor

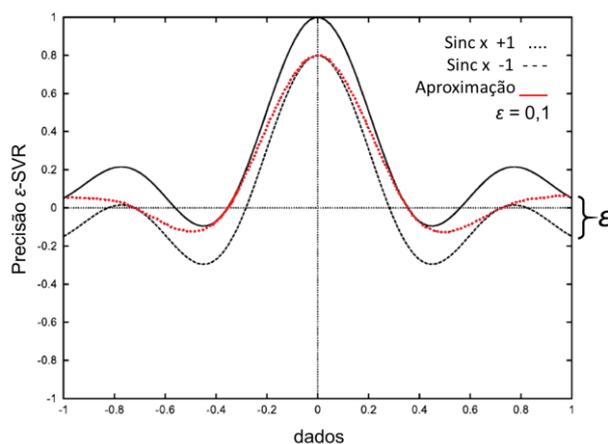
Foi observado por Cortes e Vapnik (1995) que para se construir o hiperplano ótimo é necessário uma pequena quantidade de amostras de treinamento, os vetores de suporte, nestes determina-se a margem de separação entre as classes. Os vetores de treinamento são separados sem erros por meio do hiperplano ótimo. A probabilidade de se haver um erro devido a essa “pequena” quantidade de dados de treinamento está ligada à uma razão entre o número esperado de vetores de suporte e o número de vetores de treinamento, sendo assim, os hiperplanos podem ser construídos a partir de um “pequeno” número de vetores de suporte de um conjunto de treinamento e a habilidade de generalização será alta mesmo em espaços de dimensões infinitas.

### 3.3.5 Vetores de suporte de regressão

A SVM pertence à classe de algoritmos de aprendizado supervisionado, é geralmente aplicada em problemas de classificação, mas pode também ser utilizada em problemas de regressão sendo então denominadas de vetores de suporte de regressão (do inglês *Support Vector Regression*, SVR). O conceito de SVR para estimação de função ou regressão é semelhante ao conceito da SVM. Segundo Silva Ribeiro (2012), a ideia passa por imaginar um tubo em volta da linha da função de aproximação. Como demonstrado por Scholkopf e Smola (2002) na Figura 3.10 as linhas  $\text{Sinc } x + 1$  e  $\text{Sinc } x - 1$  representam a largura do tubo- $\epsilon$ , também conhecido como  $\epsilon$ -SV regressão. A linha dentro do tubo é a função de aproximação

ou de regressão estimada, portanto, o objetivo é encontrar a função que tenha um desvio  $\varepsilon$  máximo sobre todos os valores de características  $y_i$  dos dados de treinamento e que ao mesmo tempo seja mais plana possível no espaço de alta dimensionalidade, sendo assim o erro não é notado enquanto estiver dentro da largura do tubo e nenhum desvio acima de  $\varepsilon$  é permitido.

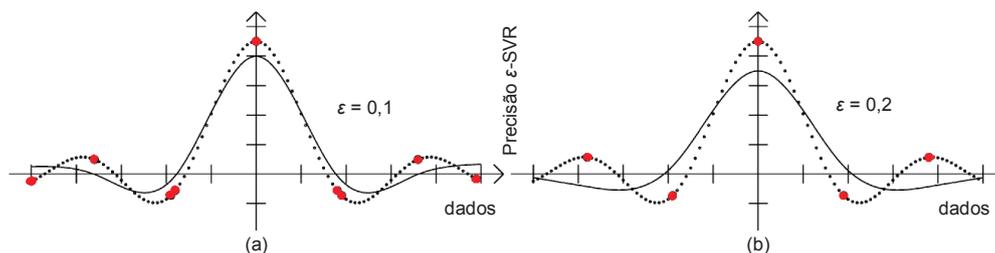
Figura 3.10 – Exemplo de estimação de função Sinc x com aproximação  $\varepsilon$ -SVR



Fonte: Adaptado de Smola e Schölkopf (2004)

Na Figura 3.11 a linha sólida é o resultado da regressão, os pontos menores são os dados de características e os maiores são os vetores de suporte para a precisão  $\varepsilon = 0,1$  e  $\varepsilon = 0,2$ , conforme apresentado nas Figuras 3.11(a) e 3.11(b).

Figura 3.11 – Vetores de suporte no  $\varepsilon$ -SVR



Fonte: Adaptado de Smola e Schölkopf (2004)

Smola e Schölkopf (2004) demonstraram a relação entre a qualidade de aproximação e a esparsividade no caso dos SVR, quanto menor a precisão requerida para a aproximação com os dados originais, menor a quantidade de vetores de suporte necessários para encontrar a função, os dados que não são vetores são redundantes conforme expressado na Figura 3.11, ou seja, mesmo

sem estes padrões os SVR seriam capazes de construir exatamente a mesma função.

### 3.3.6 O problema de regressão

Para lidar com problema de regressão faz-se necessário inicialmente apresentar o modelo de forma linear para depois obter-se uma extensão para os casos de modelos não-lineares.

A partir de um conjunto de dados de treinamento  $\{(x_1, y_1), \dots, (x_n, y_n)\} \subset \mathcal{X} \times \mathbb{R}$ , onde  $\mathcal{X}$  representa o espaço de características (por exemplo  $\mathcal{X} = \mathbb{R}^d$ ) desse sistema os quais serão distribuídos idêntica e independentemente de acordo com a função  $F(x, y) = F(x)F(y|x): \{(x_1, y_1), \dots, (x_n, y_n)\}$ .

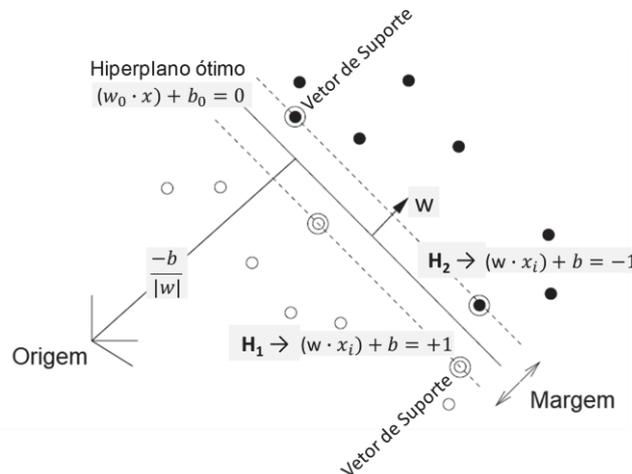
Tomando-se um problema clássico de SVM, a função linear de aproximação dada pela equação (3.16) como sendo um hiperplano:

$$f(x) = (w \cdot x) + b \quad \text{sendo } w \in \mathcal{X}, b \in \mathbb{R} \quad (3.16)$$

onde  $w$  e  $b$  são respectivamente vetor (de peso) e constante (*bias*) e conforme Vapnik et al. (1999) representam geometricamente a posição de um hiperplano.

Segundo Burges (1998),  $w$  é normal ao hiperplano,  $|b|/||w||$  é a distância perpendicular do hiperplano para a origem e  $||w||$  é a norma Euclidiana de  $w$  e  $x$  um determinado ponto a ser classificado conforme apresentado na Figura 3.12. Para Haykin (2001), maximizar a margem e separação é equivalente a minimizar a norma Euclidiana do vetor peso  $w$ .

Figura 3.12 – Separação linear de hiperplanos



FONTE: Adaptado de Burges (1998)

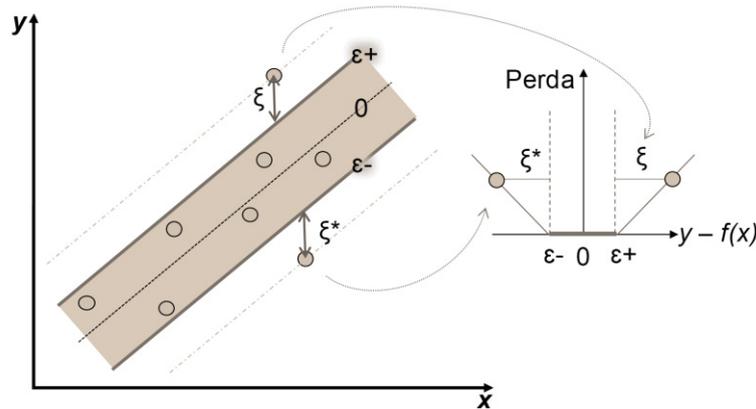
O produto escalar entre  $w$  e  $x$  determina o quão plana (*flatness*) é a função num espaço de maior dimensionalidade como por exemplo  $\mathbb{R}^3$  ou maior, ou seja, encontrar uma função que represente, que aproxime os pares  $(x_i, y_i)$  com uma precisão  $\varepsilon$ , em outras palavras se esse problema de otimização convexa é possível.

A minimização da norma Euclidiana  $w$  será determinada pela busca de seu menor valor utilizando-se a equação (3.17) que caracteriza um problema quadrático de otimização convexa, cujo principal objetivo é encontrar o hiperplano ótimo de separação entre as classes dos dados de entrada sujeito às restrições em (3.17).

$$\begin{aligned} & \text{minimização} \quad \frac{1}{2} \|w\|^2 \\ & \text{sujeito às restrições} \quad \begin{cases} y_i - (w \cdot x_i) - b \leq \varepsilon \\ (w \cdot x_i) + b - y_i \leq \varepsilon \\ \forall i = 1, \dots, n \end{cases} \end{aligned} \quad (3.17)$$

Assume-se que a função existe a partir dos dados de entrada e, portanto, será capaz de aproximar todos os pares  $(x_i, y_i)$  com uma precisão  $\varepsilon$  (Figura 3.13), sendo assim, o problema de otimização convexa torna-se viável, ou seja, a minimização da norma  $w$  estará sujeita a valores menores ou iguais ao desvio ou erro  $\varepsilon$  para a melhor aproximação por convecção possível, de cada ponto  $(x_i, y_i)$ . O erro  $\varepsilon$  é o erro que se deseja assumir.

Figura 3.13 – Margens soft, funções de aproximação  $\varepsilon$ -SVR (esquerdo) e perda  $\varepsilon$ -insensível (direito)



Fonte: Adaptado de Scholkopf e Smola (2002)

Cortes e Vapnik (1995) propuseram também uma margem *soft* devido aos casos em que não é possível separar classes sem erro e assim evitar que a função objetivo (3.17) não tenha solução. Segundo Zhou et al. (2016), a introdução das variáveis de folga não só satisfaz aos requisitos de minimização do risco estrutural da dimensão VC mas também proporcionam “boa” capacidade de generalização da função. Segundo Smola e Schölkopf (2004), analogamente à função de perda de margem *soft* utilizada em SV por Cortes e Vapnik (1995), pode-se introduzir as variáveis de folga não negativas  $\xi_i, \xi_i^* \geq 0$  (para todos  $i = 1, \dots, n$ ) em relação ao  $\varepsilon$ , conforme ilustrado na Figura 3.14, caso contrário não seria possível a separação de classes devido às restrições do processo de minimização em (3.17).

Adicionando-se as variáveis de folga ao problema (3.17) tem-se o problema de otimização, conforme Smola e Schölkopf (2004) chamado de função primal:

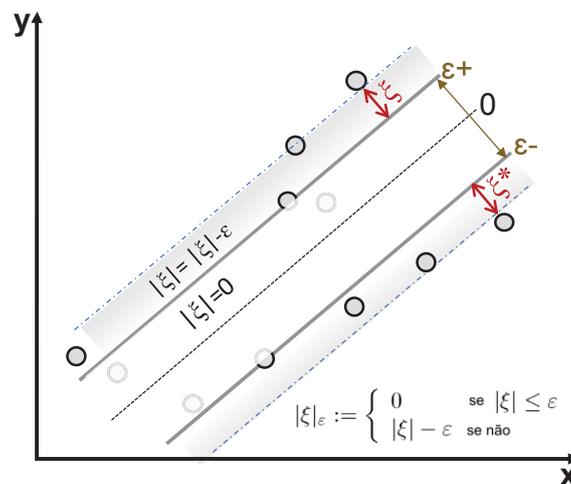
$$\begin{aligned}
 \text{minimização} \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\
 \text{sujeito às restrições} \quad & \begin{cases} y_i - (w \cdot x_i) - b \leq \varepsilon + \xi_i \\ (w \cdot x_i) + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}
 \end{aligned} \tag{3.18}$$

onde  $C > 0$  é a constante que determina a relação ou compromisso entre o quão plana será a função  $f$  e o quanto de desvio maior que  $\varepsilon$  será admitido.

Desta relação surge o conceito da função de perda chamada de  $\varepsilon$ -insensível conforme a função  $|\xi|_\varepsilon$ , apresentado na equação (3.19) e ilustrado na Figura 3.14:

$$|\xi|_\varepsilon = \begin{cases} 0 & , \text{ se } |\xi| \leq \varepsilon \\ |\xi| - \varepsilon & , \text{ se não} \end{cases} \quad (3.19)$$

Figura 3.14 – Função de perda  $\varepsilon$ -insensível com variáveis de folga  $\xi_i, \xi_i^*$



Fonte: O autor

Apenas os pontos que estão fora da região em cinza contribuem e os pontos dentro serão penalizados pela função. De acordo com Cortes e Vapnik (1995), para valores suficientemente grandes de  $C$  em (3.18), o vetor  $w$  com a constante  $b$  que minimizarão a função sob as condições (3.18) e determinarão o hiperplano que minimizará o número de erros no conjunto de treinamento e separará o resto dos elementos com a margem máxima.

Segundo Smola e Schölkopf (2004), o problema de otimização (3.18) pode ser melhor solucionado graças à sua formulação dual que é a chave para a aplicação de SVM em funções não-lineares. O método dual padrão utiliza os multiplicadores ou duais de Lagrange conforme a equação (3.20) a partir da função objetivo (3.18), dessa forma o conjunto de multiplicadores de Lagrange  $\eta_i, \eta_i^*, \alpha_i, \alpha_i^*$  funcionam como restrições à função objetivo, conforme comentado em Lorena e Carvalho (2007), tal que

$$\begin{aligned}
L = & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) - \sum_{i=1}^n (\eta_i \xi_i + \eta_i^* \xi_i^*) \\
& - \sum_{i=1}^n \alpha_i (\varepsilon + \xi_i) - y_i + (w \cdot x) + b \\
& - \sum_{i=1}^n \alpha_i^* (\varepsilon + \xi_i) + y_i - (w \cdot x) - b
\end{aligned} \tag{3.20}$$

Então para se resolver o problema da separação de dados utiliza-se os duais de *Lagrange* dados pela equação (3.20) os quais tem de satisfazer à restrição  $\alpha_i^*, \eta_i^* \geq 0$ , sendo assim, pelo princípio do ponto de sela<sup>3</sup> (*saddle point*) quando deriva-se parcialmente (3.20) em relação à  $w$ ,  $b$  e  $\xi_i^*$ , os duais terão seus valores igualados a zero e ficando o problema então regido por:

$$\frac{\partial L}{\partial b} = \sum_{i=1}^n (\alpha_i^* - \alpha_i) = 0 \tag{3.21}$$

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \tag{3.22}$$

$$\frac{\partial L}{\partial \xi_i^*} = C - \alpha_i^* - \eta_i^* = 0 \tag{3.23}$$

Após, substituindo-se (3.21), (3.22), (3.23) em (3.20) produz o problema de otimização dual:

$$\begin{aligned}
& \text{maximizar} \begin{cases} -\frac{1}{2} \sum_{i,j=1}^n (\alpha_i - \alpha_i^*) (\alpha_j + \alpha_j^*) (x_i \cdot x_j) \\ -\varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) + \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*) \end{cases} \\
& \text{sujeito à} \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \quad e \quad \alpha_i + \alpha_i^* \in [0, C]
\end{aligned} \tag{3.24}$$

---

<sup>3</sup> *Saddle point* ou ponto de sela, ponto em que uma função tem um local máxima em uma direção e local mínima em outra, sendo assim diz-se ser um ponto estável e não de inflexão <https://www.khanacademy.org/math/multivariable-calculus/applications-of-multivariable-derivatives/optimizing-multivariable-functions/a/maximums-minimums-and-saddle-points> em Maio/2018

Ao se derivar a equação (3.24) elimina-se os duais  $\eta_i, \eta_i^*$ , pela condição (3.23). Reescrevendo-se a equação (3.22) obtém-se a equação conhecida como Expansão de SV (do inglês *Support Vector Expansion*) e dada por:

$$w = \sum_{i=1}^n (\alpha_i - \alpha_i^*) x_i \rightarrow f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) (x_i \cdot x_j) + b \quad (3.25)$$

Pode-se descrever  $w$  como sendo apenas uma combinação linear dos dados de treinamento  $x_i$  de acordo com Smola e Schölkopf (2004), não dependendo da complexidade dos SV em relação a dimensão do espaço de entrada de dados mas apenas do número de vetores de suporte.

Para o cálculo de  $b$  observa-se as condições *Karush-Kuhn-Tucker* (KKT) as quais estabelecem que o produto entre as variáveis duais e suas restrições devem ser nulos:

$$\begin{aligned} \alpha_i (\varepsilon + \xi_i - y_i + (w \cdot x_i) + b) &= 0 \\ \alpha_i^* (\varepsilon + \xi_i^* + y_i - (w \cdot x_i) - b) &= 0 \end{aligned} \quad (3.26)$$

e

$$\begin{aligned} (C - \alpha_i) \xi_i &= 0 \\ (C - \alpha_i^*) \xi_i^* &= 0 \end{aligned} \quad (3.27)$$

Disto conclui-se que, somente amostras  $(x_i, y_i)$  que correspondam a  $\alpha_i^* = C$  estão fora do tubo  $\varepsilon$ -insensível e que se  $\alpha_i, \alpha_i^* = 0$  poderá nunca existir um conjunto de variáveis duais que serão simultaneamente não zeros, isto permite concluir que

$$\varepsilon - y_i + (w \cdot x_i) + b \geq 0 \quad e \quad \xi_i = 0 \quad se \quad \alpha_i < C \quad (3.28)$$

$$\varepsilon - y_i + (w \cdot x_i) + b \leq 0 \quad se \quad \alpha_i > 0 \quad (3.29)$$

Da mesma forma para  $\alpha_i^*$  tem-se:

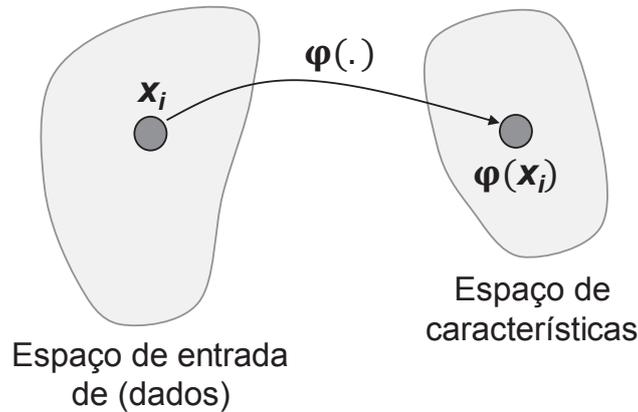
$$\begin{aligned} \max\{-\varepsilon + y_i - (w \cdot x_i) | \alpha_i < C \quad ou \quad \alpha_i^* > 0\} &\leq b \leq \\ \min\{-\varepsilon + y_i - (w \cdot x_i) | \alpha_i > 0 \quad ou \quad \alpha_i^* < C\} & \end{aligned} \quad (3.30)$$

Finalizando, da equação (3.26) conclui-se que para casos  $|f(x_i) - y_i| \geq \varepsilon$  os multiplicadores de Lagrange deverão ser diferentes de zero, em outras palavras, todas as amostras dentro do tubo- $\varepsilon$   $\alpha_i, \alpha_i^*$  sumirão como na Figura 3.14 e da mesma forma para casos  $|f(x_i) - y_i| < \varepsilon$ , a segunda parte de (3.26) é diferente de zero por isso  $\alpha_i, \alpha_i^*$  terão de ser iguais a zero para satisfazer as condições de KKT. Consequentemente tem-se uma expansão de esparsividade de  $w$  em relação a  $x_i$ , por exemplo, nem todos os  $x_i$  tem de ser descritos em  $w$ , sendo assim, os exemplos com coeficiente que não somem são chamados de vetores de suporte. Segundo Gonçalves (2009), os multiplicadores de *Lagrange*  $\alpha_i$  assumem valores positivos para exemplos de treinamento que estão a uma distância do hiperplano ótimo igual à largura da margem, ou seja, os vetores de suporte estão sob a margem. Para o restante dos exemplos de treinamento  $\alpha_i$  é nulo onde conclui-se que o hiperplano ótimo é obtido unicamente pelos “vetores de suporte”.

### 3.3.7 Funções *kernel*

“Um dos ingredientes cruciais para a SVM é o denominado *Kernel Trick* (truque de núcleo) que permite a computação de produtos escalares em espaços de características de alta dimensão utilizando funções simples, definidas em pares de padrões de entrada. Este truque permite a formulação de variantes não-lineares para qualquer algoritmo que possa ser expresso em termos de produto escalar, destes o mais promissor é o de SVM” Schölkopf (2001).

O mapeamento não-linear de um vetor de entrada para um espaço de características de alta dimensionalidade é necessário para se construir um hiperplano de separação ótimo, agora definido como uma função linear de vetores retirados do espaço de características em vez do espaço original (Figura 3.15). Em outras palavras o espaço de entrada com padrões não linearmente separáveis é transformado em espaço de características de alta dimensionalidade onde os padrões passam a ser linearmente separáveis. Para Suykens et al. (2002), a aplicação do *Kernel trick* permite trabalhar com espaços de características de dimensões enormes sem explicitamente ter que realizar o processamento neste espaço, mas sim em outro espaço pela utilização do *kernel trick*, transformando o problema primal e um problema dual.

Figura 3.15 – Mapeamento não-linear  $\varphi_i$  do espaço de entrada para o de características

Fonte: Adaptado de Haykin (2001)

Segundo Lima (2004), as SVMs abrem a possibilidade de se operar com novas representações eficientes de funções não-lineares, empregadas no projeto de algoritmos de aprendizado, estas representações fazem uso das chamadas funções *kernel*.

Conforme Kung (2014), dentre as funções *kernel* mais utilizadas estão as gaussianas ou funções de base radial (do inglês *Radial Basis Function* – RBF), análise de componentes principais (do inglês *Principal Component Analysis* – PCA) e as polinomiais conforme ilustrado no Quadro 3.2.

Quadro 3.2 – Funções *kernel* aplicáveis a SVM

Tipo	Função $k(x \cdot x')$	Parâmetros
Polinomial	$(\delta(x \cdot x') + k)^d$	$\delta, k, d$
Gaussiano	$\exp(-\sigma \ x \cdot x'\ ^2)$	$\sigma$
Sigmoide	$\tanh(\delta(x \cdot x') + k)$	$\delta, k$

Adaptado de Lorena e Carvalho (2007)

Para se obter vetores de suporte a partir de padrões não-lineares basta fazer o pré-processamento dos padrões de treinamento  $x_i$  pelo mapa  $\Phi: \mathcal{X} \rightarrow \mathcal{H}$ , sendo  $\mathcal{H}$  um espaço de características conforme dito por Smola e Schölkopf (2004). Considerando-se o mapa  $\Phi: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  sendo  $\Phi = (x_1^2, \sqrt{2}x_1x_2, x_2^2)$  entende-se que para este caso  $x \in \mathbb{R}^2$ , então o pré-processamento da SVM produzirá uma função quadrática. Se realizarmos uma expansão, essa abordagem para um caso de

características polinomiais de ordens maiores e de alta dimensionalidade o uso da SVM pode se tornar inviável, porém se o uso de um *kernel* for introduzido o problema passará a ser possível de ser executado.

Segundo Schölkopf (2001), a noção de similaridade dentro de espaços euclidianos baseia-se em conceitos geométricos utilizados em SVM, isso quer dizer ângulos e comprimentos em algum espaço de alta dimensionalidade. Para se evitar trabalhar com espaços de altas dimensionalidades tenta-se se pegar um espaço de características no qual o produto escalar possa ser validado diretamente utilizando uma função não-linear no espaço de entrada por meio do *kernel trick* em (3.31), tal que

$$k(x, x') = \langle \Phi(x), \Phi(x') \rangle \quad (3.31)$$

Segundo Haykin (2001), pode-se usar o produto interno do *kernel*  $K(x, x')$  para construir o hiperplano ótimo no espaço de características sem ter que considerar o próprio espaço de forma explícita.

Partindo-se de Boser et al. (1992), tem-se do mapa de características  $\Phi: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ :

$$\langle (x_1^2, \sqrt{2}x_1x_2, x_2^2), (x_1'^2, \sqrt{2}x_1'x_2', x_2'^2) \rangle = \langle x, x' \rangle^2 \quad (3.32)$$

Como descrito anteriormente os vetores de suporte dependem exclusivamente do produto escalar entre os diversos padrões dos dados de entrada, desta forma se o algoritmo do SV depende apenas do produto escalar entre os padrões de entrada  $x_i$  então é suficiente apenas conhecer (3.31) do que  $\Phi$  explicitamente, permite-se então reescrever (3.24) com o *kernel*:

$$\begin{aligned} & \text{maximizar} \begin{cases} -\frac{1}{2} \sum_{i,j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)k(x_i, x_j) \\ -\varepsilon \sum_{i=1}^n (\alpha_i + \alpha_i^*) + \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*) \end{cases} \quad (3.33) \\ & \text{sujeito a} \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0 \quad e \quad \alpha_i, \alpha_i^* \in [0, C] \end{aligned}$$

Da mesma forma aplicando-se o mapa em (3.25) tem-se

$$w = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \Phi x_i \rightarrow f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) k(x_i, x_j) + b \quad (3.34)$$

A diferença para o caso linear é que  $w$  não é mais dado explicitamente e que em casos não-lineares o problema de otimização corresponde a encontrar a superfície mais plana possível no espaço de características e não no espaço de entrada.

Conforme Smola e Schölkopf (2004), para se utilizar uma função *kernel* precisa-se conhecer quais  $k(x, x')$  correspondem ao produto escalar no espaço de características  $\mathcal{H}$ , portanto algumas condições têm de serem satisfeitas. Existe uma dedução de tais condições baseadas no Teorema de Mercer (1909), como combinação linear positiva de *kernels*, integral de *kernels* além do próprio produto escalar entre eles, que não serão desenvolvidas aqui. Segundo Haykin (2001) o teorema de Mercer diz se um *kernel* candidato é realmente um *kernel* de produto escalar em algum espaço, portanto admissível em uma SVM ou não, entretanto ele não trata de como construir tais funções.

O significado do teorema de Mercer de maneira menos formal, segundo Smola e Schölkopf (2004):

$$\int_{x \times x} k(x, x') f(x) f(x') dx d' \geq 0 \quad \text{para todos } f \in L_2(\mathcal{X}) \quad (3.35)$$

Se (3.35) se mantem, pode-se escrever  $k(x, x')$  como um produto escalar em um espaço de características, destas condições conclui-se que algumas regras simples para a composição de funções *kernel*, as quais satisfaçam também as condições de Mercer.

### 3.3.8 Método SVM por mínimos quadrados

Segundo Suykens et al. (2002), a formulação original de Vapnik para SVM enfatiza interpretações primal-dual e faz uso do *Kernel trick*, este pode ser aplicado de maneiras diferentes, geralmente técnicas lineares utilizadas na resolução de padrões não-lineares. O método SVM torna-se mais fácil de se formular em termos de mínimos quadrados (do inglês *Least Squares SVM* – LSSVM) e de restrições por igualdade (*equality*) do que outras funções de perda ou restrições por desigualdade (*inequality*), portanto LSSVM é uma modificação na formulação do classificador SVM que busca a solução baseada em um sistema linear ao invés do problema de programação quadrática.

De maneira similar a formulação clássica para SVM em LSSVM tem-se no espaço primal (3.36) em (3.37) tal que

$$y(x) = \text{sign}[w^T \varphi(x) + b] \quad (3.36)$$

onde  $w$  é vetor de pesos na função primal,  $b$  é a constante de polarização (*bias*) e  $\varphi(\cdot)$  é a função de mapeamento no espaço de alta dimensionalidade.

Como definido por Suykens et al. (2002), a formulação de Vapnik é modificada em dois pontos: primeiro, ao invés de restrições por desigualdade toma-se restrições por igualdade como em (3.37):

$$\begin{aligned} \min_{w, b, e} \quad J_P(w, e) &= \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{k=1}^n e_k^2 \\ \text{sujeito às restrições} \quad y_k [w^T \varphi(x_k) + b] &= 1 - e_k \quad k = 1, \dots, n \end{aligned} \quad (3.37)$$

onde  $J$  é a função de custo que consiste da soma do erro quadrático de ajuste e um termo de regularização,  $\gamma$  o parâmetro de regularização<sup>4</sup> e  $e_k$  é a variável de erro inserida em LSSVM.

Como apresentado em (3.37), o valor 1 do lado direito da formula de restrição é considerado mais como um valor a ser alcançado e não como um

---

<sup>4</sup> Também apresentado como “C” em (3.18) página 51.

limitador, sobre este valor (propósito) uma variável de erro é permitida de tal forma que erros de classificação para os casos de sobreposições de distribuições possam ser tolerados. Esta variável de erro funciona de forma similar as variáveis de folga da SVM. Na segunda modificação, a função de perda quadrada é utilizada pela variável de erro  $e_k$ , estas modificações simplificarão o problema. Para maior detalhamento da formulação, este pode ser encontrado em Suykens et al. (2002).

O dual da LSSVM expresso por (3.38) é obtido de forma semelhante ao dual da SVM, por derivadas e dos multiplicadores de Lagrange, sendo possível expressar a solução em termos de variáveis de erro desconhecidas 'e' pela eliminação dos multiplicadores de Lagrange  $\alpha_k$  ao invés do erro e.

$$y(x) = \text{sign} \left[ \sum_{k=1}^n \alpha_k y_k K(x, x_k) + b \right] \quad (3.38)$$

Uma das propriedades da LSSVM é a possibilidade do uso de parâmetros de ajuste (do inglês *tuning*) permitindo encontrar a melhor performance do conjunto de dados de validação sendo  $\gamma$  é o parâmetro de regularização e a largura da banda da *kernel*. Conforme De Brabanter et al. (2013), regressão não paramétrica é uma ferramenta de análise de dados sendo flexível para modelagem não-linear, dessa forma StatLSSVM (*Statistical LSSVM*) emprega a técnica aliada à inferência estatística na LSSVM via ferramentas do ambiente computacional Matlab®, onde

$$\overbrace{\min_{w, b, e}} J_P(w, e) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{k=1}^n v_k e_k^2 \quad (3.39)$$

$$\text{sujeito às restrições} \quad y_k = [w^T \varphi(x_k) + b + e_k] \quad k = 1, \dots, n$$

onde  $v_k$  representa o peso do  $k$ -ésimo residual.

A StatLSSVM procura tornar a função de perda robusta o que levaria ao problema de programação quadrática então optou-se por um método simplificado no qual resolve-se o problema LS por pesos em cada iteração até que um certo critério de parada seja satisfeito, a partir da formulação LSSVM em (3.37) chega-se a formulação de StatLSSVM (3.39), detalhamento por ser encontrado em De Brabanter et al. (2013).

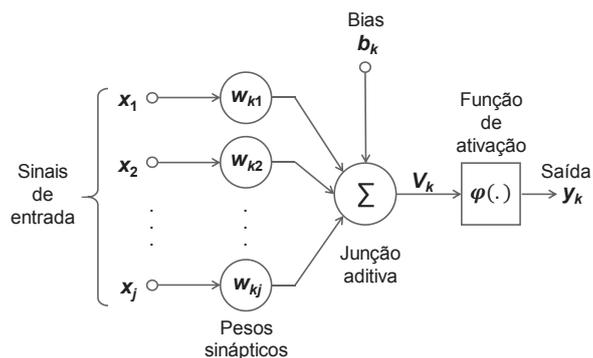
### 3.4 REDES NEURAIS ARTIFICIAIS

Segundo Haykin (2001), “Uma rede neural artificial (do inglês *Artificial Neural Network* – ANN) é um processador paralelamente distribuído, constituído de unidade de processamento simples, que tem a propensão natural para armazenar conhecimento e torná-lo disponível para o uso. Ela se assemelha ao cérebro em dois aspectos: o conhecimento é adquirido pela rede a partir de seu ambiente por meio de um processo de aprendizado e pelas forças (intensidades) de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido”.

#### 3.4.1 Neurônio artificial

A primeira ideia de um neurônio artificial veio da comparação com o neurônio biológico do trabalho pioneiro de McCulloch-Pitts em 1943, o qual descrevia um neurônio artificial com entradas produzindo saídas baseadas em pesos fixos como descrito por Haykin (2001). A evolução do neurônio de McCulloch-Pitts foi o *perceptron* proposto por Rosenblatt em 1958, pois, seus pesos passaram a ser ajustáveis possibilitando o aprendizado, porém limitado a resolução de problemas linearmente separáveis. Conforme Haykin (2001), esse foi um método inovador de aprendizado supervisionado em que o neurônio (*perceptron*) tem a capacidade de assumir em sua saída o valor 1 se o campo local induzido é não-negativo e 0 caso contrário, o neurônio produz um conjunto de sinapses cada uma caracterizada como sendo um peso  $w_{kj}$  onde é conectado um sinal de entrada  $x_j$  (Figura 3.16). Os sinais são somados, ponderados pelas respectivas sinapses do neurônio constituindo uma combinação linear. Uma função conhecida como de ativação, também é utilizada para restringir a amplitude na saída do neurônio, a amplitude é um intervalo que restringe a saída geralmente normalizado entre [0,1] ou [-1,1].

Figura 3.16 – Neurônio artificial de McCulloch-Pitts



Fonte: Adaptado de Haykin (2001)

O modelo na Figura 3.16 tem também uma variável de bias<sup>5</sup>  $b_k$  como sendo um valor externo e tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação.

Um neurônio pode ser descrito matematicamente como

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (3.40)$$

e

$$y_k = \varphi(u_k + b_k) \quad (3.41)$$

onde  $k$  é o neurônio e  $j$  a sinapse conectada a ele, assim  $x_j$  significa uma entrada na sinapse  $j$  conectada ao neurônio  $k$  e multiplicada pelo peso sináptico  $w_{kj}$ ,  $u_k$  a saída do combinador linear e  $b_k$  o *bias* aplicado em (3.41).

A função de ativação é dada por  $\varphi$  por fim  $y_k$  a saída do neurônio tal que

$$v_k + u_k + b_k \quad (3.42)$$

onde  $v_k$  é o potencial de ativação do neurônio  $k$ .

---

<sup>5</sup> *Bias* em ANN possui significado diferente dos definidos anteriormente para LSN e SVM conforme citado na nota de rodapé 2.

Combinando-se as equações (3.40, 3.41 e 3.42) tem-se

$$v_k = \sum_{j=1}^m w_{kj} x_j \quad (3.43)$$

$$y_k = \varphi(v_k) \quad (3.44)$$

Sendo assim, adicionando-se uma nova sinapse, tem-se a entrada

$$x_0 = +1 \quad (3.45)$$

e o seu peso

$$w_{k0} = b_k \quad (3.46)$$

### 3.4.2 Funções de ativação

Conforme Haykin (2001), são identificadas três tipos básicos de funções de ativação: a limiar, a linear por partes e a função sigmoide, definidas como a seguir (3.47) a (3.50):

$$\text{Limiar} \quad \varphi(v) = \begin{cases} 1 & \text{se } v \geq 0 \\ 0 & \text{se } v < 0 \end{cases} \quad (3.47)$$

$$\text{Linear} \quad \varphi(v) = v \quad (3.48)$$

$$\text{Saturação} \quad \varphi(v) = \begin{cases} +1 & \text{se } v > 1 \\ v & \text{se } v \in [-1, 1] \\ -1 & \text{se } v < -1 \end{cases} \quad (3.49)$$

$$\text{Sigmoide} \quad \varphi(v) = \frac{1}{1 + \exp(-av)} \quad (3.50)$$

onde  $a$  indica a declividade da função sigmoide

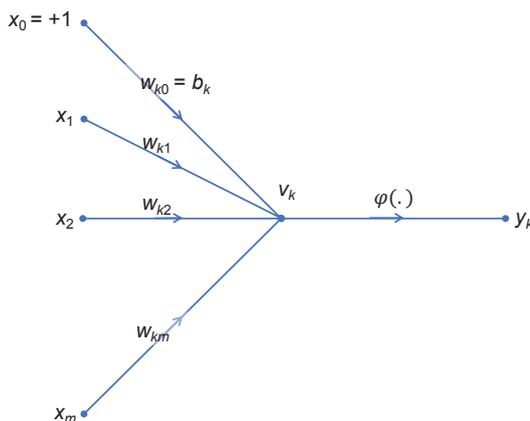
Em alguns casos a função limiar (3.47) pode assumir valores entre -1 e 1 quando existir uma função de ativação ímpar do campo local induzido, sendo +1 se  $v > 0$ , 0 se  $v = 0$  e -1 se  $v < 0$ . Normalmente conhecida como função sinal cuja forma é de uma função tangente hiperbólica definida por:

$$\varphi(v) = \tanh(v) \quad (3.51)$$

### 3.4.3 Grafos orientados

Segundo Haykin (2001), uma rede neural é um grafo orientado da Figura 3.16 constituído de nós com elos de interligação sinápticos e de ativação como apresentado na forma simplificada na Figura 3.17:

Figura 3.17 – Grafo de fluxo de sinal de um neurônio



Adaptado de Haykin (2001)

cujas propriedades são:

- Cada neurônio é representado por um conjunto de elos sinápticos lineares, um *bias* aplicado externamente e um elo de ativação possivelmente não-linear. O *bias* é representado por um elo sináptico conectado a uma entrada fixa +1.
- Os elos sinápticos de um neurônio ponderam os seus respectivos sinais de entrada.

- A soma ponderada dos sinais de entrada define o local induzido do neurônio em questão.
- O elo de ativação limita o campo local induzido do neurônio para produzir uma saída, sendo o estado do neurônio definido em termos de seu campo local induzido ou de sua saída.

Um grafo orientado é caracterizado por um nó de controle e um nó computacional além de elos de comunicação sendo que o nó de fonte fornece sinal de entrada, o nó computacional representa um único neurônio e os elos de comunicação são a conexão entre os nós de fonte e os nós computacionais.

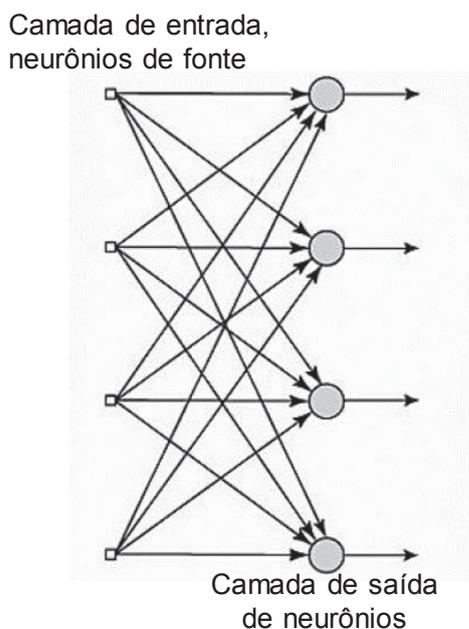
#### 3.4.4 Arquiteturas de rede

Arquiteturas (estruturas) de rede são a maneira pela qual os neurônios de uma rede neural estão estruturados, está intimamente ligada com o algoritmo (regras) de aprendizagem usado para treinar a rede, são apresentadas em três classes distintas:

- 1) Redes alimentadas adiante com camada única;
- 2) Redes alimentadas diretamente com múltiplas camadas;
- 3) Redes recorrentes.

Em uma rede neural em camadas, os neurônios estão organizados na forma de camadas, mas não vice-versa, portanto a rede é estritamente alimentada adiante ou acíclica, conforme ilustrado na Figura 3.18.

Figura 3.18 – Rede alimentada adiante, uma camada de neurônios

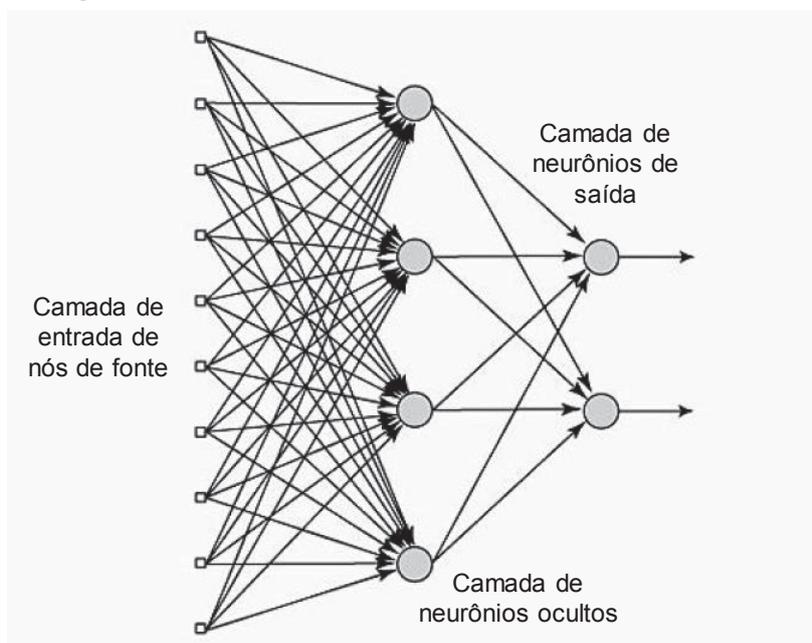


Fonte: Adaptado de Haykin (2001)

As redes neurais artificiais constituídas de uma ou mais camadas ocultas são conhecidas como *perceptron* multicamadas (do inglês *Multi-layer Perceptron* – MLP). Conforme apresentado na Figura 3.19, as camadas possuem “n” neurônios totalmente conectados. A MLP possui uma camada de entrada de nós de fonte a qual fornece o sinal para os neurônios ocultos ou unidades ocultas, ou seja, os nós computacionais da segunda camada (primeira camada oculta).

Os sinais de saída da segunda camada servirão de entradas para a terceira camada a qual fornecerá entrada para a próxima camada e assim por diante. Quanto maior o número de camadas a rede se torna capaz de extrair informações estatísticas de ordem elevada, essa característica é relevante quando se tem um número “acentuado” de entradas.

Figura 3.19 – Rede com uma camada oculta e uma de saída

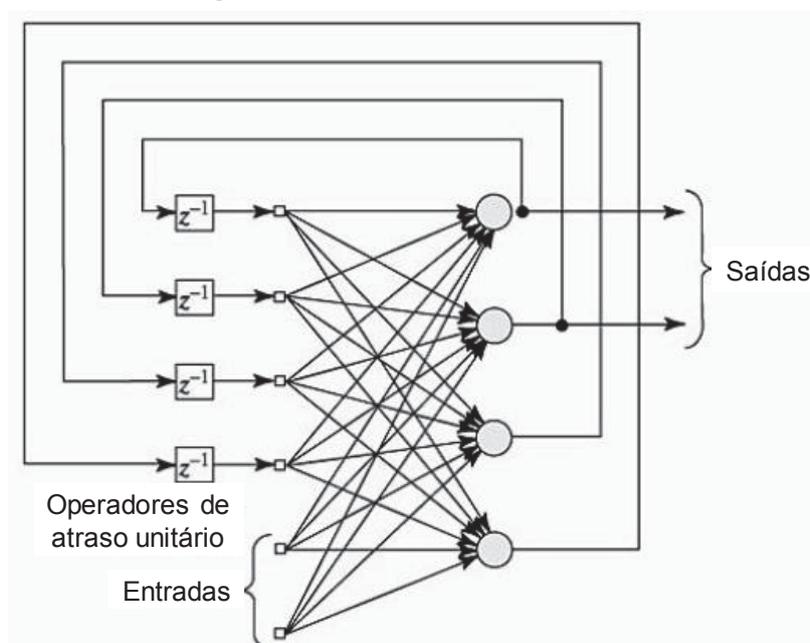


Fonte: Adaptado de Haykin (2001)

A melhora no processo de aprendizagem da rede ocorre a cada iteração por meio do ajuste dos pesos sinápticos e pelos níveis de polarização adaptando-se ao meio no qual estiver inserida, sendo assim, uma rede neural é estimulada pelo ambiente, sofre modificações nos seus parâmetros devido ao estímulo e responde de uma maneira nova ao ambiente devido as modificações ocorridas na sua estrutura interna.

As redes neurais artificiais recorrentes possuem pelo menos um laço de realimentação em relação a rede alimentada adiante (Figura 3.20), tem um impacto profundo na capacidade de aprendizado da rede e no seu desempenho. Os laços de realimentação envolvem o uso de ramos particulares compostos de elementos de atraso unitário  $z^{-1}$ , que resultam em um comportamento dinâmico não-linear caso a rede contenha unidades não-lineares.

Figura 3.20 – Rede neural recorrente



Fonte: Adaptado de Haykin (2001)

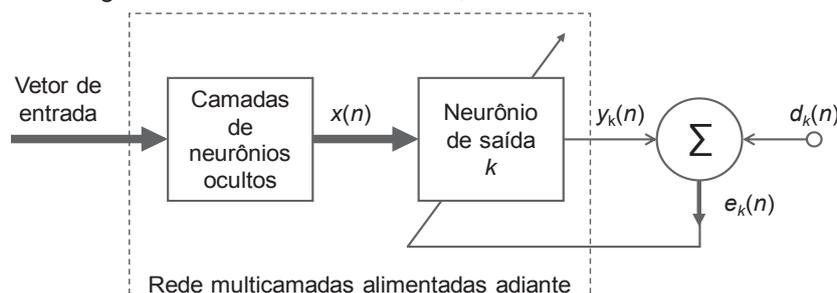
### 3.4.5 Processos de aprendizado

Aprendizado com um professor ou aprendizado supervisionado envolve a modificação de pesos sinápticos de uma rede neural pela aplicação de um conjunto de amostras de treinamento ou exemplos de tarefa, cada exemplo consiste em um sinal de entrada único e de resposta desejada correspondente. Apresenta-se à rede um exemplo escolhido ao acaso do conjunto, os pesos sinápticos da rede (parâmetros livres) são modificados para minimizar a diferença entre a resposta desejada e a resposta real da rede produzida pelo sinal de entrada.

Aprendizado (ou aprendizagem) é um processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação pelo ambiente no qual a rede está inserida. O tipo de aprendizagem é determinado pela maneira que a modificação dos parâmetros ocorre. Haykin (2001).

Desta forma, o processo de aprendizado implica numa sequência de eventos em que a ANN é estimulada por um ambiente que provoca alterações nos seus parâmetros livres, esta por sua vez responde de maneira nova ao ambiente devido a modificações em sua estrutura interna.

Figura 3.21 – Diagrama de blocos de uma ANN, um neurônio na camada de saída e erro



Fonte: Adaptado de Haykin (2001)

Na Figura 3.21 tem-se o neurônio  $k$  como parte de um nó computacional único da camada de saída de uma rede neural alimentada adiante. O vetor de saída  $x(n)$  produzido pelas camadas de neurônios ocultos aciona o neurônio  $k$  que por sua vez produz um sinal  $y_k(n)$  que não é exatamente o valor de saída desejado  $d_k(n)$ , portanto existe uma diferença entre as saídas desejada e estimada, essa diferença pode ser chamada de erro  $e_k(n)$ , expresso pela equação (3.52):

$$e_k(n) = d_k(n) - y_k(n) \quad (3.52)$$

Segundo Haykin (2001), o sinal de erro  $e_k(n)$  aciona um mecanismo de controle, cujo propósito é aplicar uma sequência de ajustes corretivos aos pesos sinápticos do neurônio  $k$  de modo a aproximar passo a passo o sinal de saída  $y_k(n)$  da resposta desejada  $d_k(n)$ , esse objetivo é alcançado pela minimização da função de custo ou índice de desempenho  $E(n)$  a partir do erro  $e_k(n)$  na equação (3.53):

$$E(n) = \frac{1}{2} e_k^2(n) \quad (3.53)$$

Este processo de aprendizado é denominado de aprendizado por correção de erro.

Outro processo de aprendizado conhecido é o baseado em memória, onde informações passadas são armazenadas em uma grande memória de exemplos de entrada-saída classificados corretamente. Em uma classificação de padrões com duas classes (classificação binária) ou hipóteses são consideradas, assim a

resposta desejada assume o valor 0 ou -1 para uma das classes e valor 1 para a segunda classe.

Para Haykin (2001), todos os algoritmos de aprendizado baseados em memória envolvem dois ingredientes essenciais: o critério de definição de vizinhança local do vetor de teste e a regra de aprendizagem aplicada aos exemplos de treinamento na vizinhança local do vetor de teste. O que difere estes algoritmos é a forma como estes dois ingredientes são definidos.

Um tipo simples de aprendizado baseada em memória é o da regra do vizinho mais próximo onde o exemplo de treinamento mais próximo do vetor de teste é determinado pela distância euclidiana entre o vetor de entrada e o vetor de teste, a classe mais próxima é apresentada como a classificação do vetor de teste.

### 3.5 CRITÉRIOS DE DESEMPENHO

Como mencionado anteriormente um modelo gera um valor estimado o qual pretende-se que seja o mais próximo dos dados de um sistema a ser modelado. Devido à complexidade do processo de modelagem de um sistema e por vezes objetivos conflitantes como nos algoritmos de otimização dos SVR que buscam minimizar os erros e maximizar a margem de busca sob certas restrições, torna-se necessário o emprego de índices de desempenho para se comparar os resultados e por meio destes, poder realizar os ajustes nos parâmetros de modo a refinar a modelagem realizada.

Dois índices serão utilizados para estes objetivos nesta dissertação:

- a) O erro quadrado médio (*Mean Squared Error*, MSE) – garante a qualidade dos estimadores ou previsores; é sempre não-negativo e valores quanto mais próximos de zero melhor.
- b) O coeficiente de correlação múltipla  $R^2$  – permite determinar quanto um modelo é capaz de representar os dados observados. Quanto mais próximo de 1 melhor o resultado, sendo 1 a indicação exata de adequação do modelo aos dados medidos.

O cálculo de  $R^2$  é dado pelo cálculo soma dos quadrados devido a regressão (*Sum of Squares to regression*, SS), a soma total dos quadrados (*Total Sum of Squares*, TSS) além da Soma residual dos quadrados (*Residual Sum of Squares*, RSS) tal que.

$$TSS = \sum_{k=1}^N \{y(k) - \bar{y}\}^2 \quad (3.54)$$

$$SS = \sum_{k=1}^N \{\hat{y}(k) - \bar{y}\}^2 \quad (3.55)$$

$$RSS = \sum_{k=1}^N \{\hat{y}(k) - y(k)\}^2 \quad (3.56)$$

$$R^2 = \frac{SS}{TSS} = 1 - \frac{RSS}{TSS} \quad (3.57)$$

O MSE mede a média dos quadrados dos erros ou desvios, ou seja, a diferença entre o previsor e o valor estimado. O MSE é dado por:

$$MSE = \frac{1}{N} \sum_{k=1}^N (y_k - \hat{y}_k)^2 \quad (3.58)$$

onde  $y_k$  é o vetor de n previsões do vetor  $\hat{y}_k$  é o vetor observado de valores da variável que está sendo prevista. O objetivo é se obter o menor valor para o MSE.

## 4 PREPARAÇÃO DE DADOS E AJUSTES DAS TÉCNICAS

Esta dissertação restringe-se ao motor de combustão a Diesel montado em um veículo de referência. As técnicas de modelagem utilizarão os dados de motor coletados com o veículo parado e em movimento em diversos ciclos de trabalho.

### 4.1 AQUISIÇÃO DOS DADOS

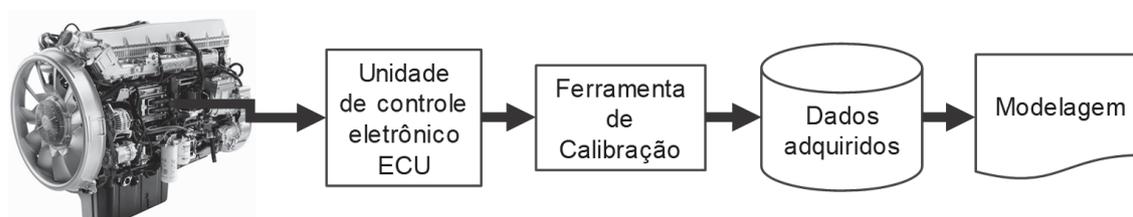
Os dados obtidos foram adquiridos em medições reais em diversos trajetos, como descritos no Capítulo 4.4, realizados por um veículo de referência reproduzindo os ciclos de trabalho a que este será submetido quando liberado para a produção e venda no mercado nacional e internacional.

Segundo Aguirre (2015), uma das etapas no processo de identificação de sistemas é a de teste e coleta de dados. Nessa etapa a modelagem do sistema é feita a partir de dados gerados de um sistema, pois, nem sempre é possível desenvolver um modelo matemático a partir das equações que regem a física do sistema. Esses dados são extraídos a partir de testes específicos que permitam representar a dinâmica do sistema. Dentre os problemas desse processo tem a escolha dos sinais de excitação, execução de testes e a escolha do período de amostragem.

Os sinais foram adquiridos via ferramenta de engenharia específica para leitura e calibração de todos os parâmetros de *software* do motor, cuja unidade de controle eletrônico (*Electronic Control Unit* – ECU) foi desenvolvida para gerenciar todo o funcionamento do motor. A taxa de aquisição ajustada na ferramenta de leitura dos parâmetros do *software* foi de 100 ms para todas as medições, essa taxa foi escolhida de maneira que todos os dados necessários para o modelo pudessem ser coletados durante todo o trajeto. A ferramenta de engenharia utilizada combina sua capacidade de processamento e armazenamento com a quantidade de sinais a serem gravados e suas respectivas taxas de aquisição. Os requisitos para aquisição de dados definidos por Nyquist estão garantidos pelo *software* da ECU e de sua capacidade de processamento, dessa forma não serão detalhados nesta dissertação.

Todas as medições foram convertidas para o formato do Matlab® com o auxílio da própria ferramenta de engenharias utilizada na aquisição dos dados.

Figura 4.1 – Diagrama de trabalho dos dados adquiridos



Fonte: O autor (2018)

## 4.2 ESCOLHA DOS DADOS DE ENTRADA E SAÍDA DO MODELO

A escolha das variáveis de entrada do sistema foi realizada com base no conhecimento prévio do comportamento do sistema de arrefecimento em projetos anteriormente lançados pela empresa produtora dos caminhões. A arquitetura dos modelos para os sinais de rotação e temperatura foi de múltiplas entradas-saída única (do inglês *Multiple Inputs-Single Output* – MISO). Para cada uma das técnicas adotadas para modelagem procurou-se a combinação entre a funcionalidade do sistema com o uso de dados reais para se treinar e validar os modelos.

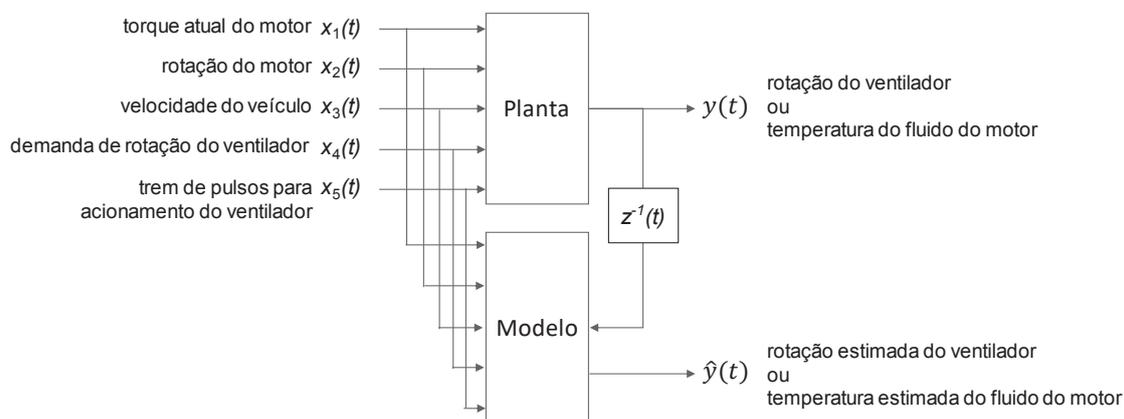
Decidiu-se pela modelagem dos sinais de rotação do ventilador e da temperatura após testes com veículo parado. Em seguida buscou-se identificar em quais momentos o ventilador foi acionado e posteriormente desacoplado. Ao se monitorar o conjunto de variáveis relacionadas ao sistema pode-se compará-las e após rodar várias combinações destas variáveis com o algoritmo LSM, pode-se observar qual combinação apresentou melhor resultado em termos de MSE e  $R^2$  conforme apresentado na Tabela 4.1 e ilustradas na Figura 4.2, onde o termo “planta” significa, no contexto desta dissertação, o motor do veículo de referência.

Tabela 4.1 – Variáveis adquiridas durante as medições

modelos	descrição da variável	unidade física
entrada	torque atual do motor	Nm
entrada	rotação do motor	rpm
entrada	velocidade do veículo	km/h
entrada	demanda calculada de rotação do ventilador	%
entrada	trem de pulsos para acionamento do ventilador	%
saída modelo 1	rotação do ventilador	rpm
saída modelo 2	temperatura do fluido do motor	°C

Fonte: O autor (2018)

Figura 4.2 – Diagrama MISO de entradas e saída dos modelos



Fonte: O autor (2018)

A configuração MISO do modelo apresentado na Figura 4.2 é válida para os dois sinais a serem modelados: rotação do ventilador e temperatura do fluido de arrefecimento do motor. Em todas as técnicas, as saídas dos modelos foram deslocadas em  $(t-1)$  em relação as suas entradas.

#### 4.2.1 Correlação dos dados dos modelos

Apesar da escolha dos sinais de entrada dos modelos apresentados na Tabela 4.1, buscou-se a correlação destes sinais com a saída dos modelos de modo a se observar quais possuíam maior importância para o mesmo. O coeficiente

de correlação utilizado foi o linear de Pearson, disponível no Matlab®, este retornou quais entradas possuem maior e menor relação com a saída dos modelos.

Para o modelo de rotação do ventilador, os sinais que se mostraram mais relevantes foram a demanda de atuação e o trem de pulsos para o acionamento do ventilador. O torque do motor foi o sinal de menor importância para o modelo, como apresentados na Tabela 4.2.

Tabela 4.2 – Correlação das entradas/saída do modelo LSM para rotação do ventilador

relevância	correlação com $y(t-1)$	entradas	$R^2$ de validação		
			1 a 5	1 a 4	1 a 3
1	0,4955	demanda calculada de rotação do ventilador			
2	0,3808	trem de pulsos para acionamento do ventilador			
3	0,2484	rotação do motor			0,4021
4	0,0990	velocidade do veículo		0,3949	
5	0,0695	torque atual do motor	0,3940		

Fonte: O autor (2018)

Em comparação ao  $R^2$  de validação com todas as entradas e sem as entradas de menor relevância pode-se constatar que os valores de desempenho não sofreram mudanças em relação ao resultado com todas as entradas, todos ficaram em torno de 39%, mostrando que o modelo poderia ser simplificado com a retirada de tais sinais.

Para o modelo de temperatura do fluido, os sinais não apresentaram boa correlação, sendo a demanda de atuação e o trem de pulsos para o acionamento do ventilador as entradas de menor influência na a saída do modelo, como apresentado na Tabela 4.3. Os índices de desempenho retirando-se as entradas de menor importância não produziram mudanças significativas em termos de desempenho, gerando  $R^2$  em torno de -0,58

Tabela 4.3 – Correlação das entradas/saída do modelo LSM para temperatura do fluido

relevância	correlação com $y(t-1)$	entradas	$R^2$ de validação		
			1 a 5	1 a 4	1 a 3
1	-0,2706	velocidade do veículo			
2	-0,2135	rotação do motor			
3	-0,0354	torque atual do motor			-0,5976
4	-0,0222	demanda calculada de rotação do ventilador		-0,5572	
5	-0,0180	trem de pulsos para acionamento do ventilador	-0,5857		

Fonte: O autor (2018)

### 4.3 NORMALIZAÇÃO DOS DADOS

Por questões de confidencialidade optou-se pela normalização dos dados adquiridos no veículo de referência. Após a conversão dos dados adquiridos para o formato adequado ao Matlab®, os dados foram normalizados utilizando-se a equação (4.1) de modo que todos os sinais adquiridos estivessem no intervalo entre 0 e 1. Assim adotou-se a equação

$$V_{normal} = \frac{V_{real} - \min(V_{real})}{\max(V_{real}) - \min(V_{real})} \quad (4.1)$$

onde  $V_{real}$  refere-se ao valor original de cada variável e  $V_{normal}$  é o resultado normalizado para cada uma das variáveis utilizadas no procedimento de modelagem do sistema de arrefecimento, conforme mencionado na Tabela 4.1..

### 4.4 ESCOLHA DO TRAJETO REALIZADO PELO VEÍCULO DE REFERÊNCIA

Inicialmente foram realizados testes aleatórios com o veículo parado para se ter um contato inicial com o sistema de arrefecimento e seu funcionamento básico, bem como o reconhecimento do veículo, motor em teste e o sistema de coleta dos dados.

Numa segunda etapa realizou-se medições em pequenos trechos dentro da empresa de modo a se entender o funcionamento do sistema de arrefecimento do motor e os momentos de demanda pelo funcionamento do ventilador e assim

decidir quais variáveis seriam utilizadas como entradas e saída dos modelos e adquiridas durante todos os trajetos medidos.

Na terceira e última etapa os dados foram obtidos a partir dos veículos que rodam em testes diários de campo, para a validação destes carregados em sua carga nominal. Os trechos realizados foram feitos em rodovias intermunicipais buscando diferentes tipos de relevo, subidas e descidas de serra, condições em altitude e ao nível do mar, entre outros. Estes testes levam também em conta as condições ambientais de temperatura, umidade e pressão atmosférica, condições estas vitais para o processo de calibração do desempenho dos veículos e do motor a combustão, sendo assim, o tamanho dos dados de entrada variou conforme o trecho percorrido pelo veículo. Como visto na Tabela 4.4 o trajeto número 1 foi escolhido para se gerar os modelos e todos os resultados apresentados nesta dissertação, este trajeto foi composto de quatro medições para o mesmo trecho percorrido. Por motivo de confidencialidade optou-se por não mencionar as localizações de cada um dos trajetos realizados pela empresa produtora dos caminhões em seus testes dentro de projetos de desenvolvimento.

Para se garantir uma distribuição homogênea nos dados destes trechos, tanto na fase de treinamento como validação do modelo, optou-se por realizar uma distribuição aleatória das amostras coletadas antes de se realizar o processo de modelagem. Enfatizando que o foco foi a obtenção de modelos baseados em regressão e não de identificação de sistemas, onde a dinâmica temporal é relevante.

Tabela 4.4 – Medições por trajetos x amostras

<b>trajetos</b>	<b>número total de amostras por trajeto</b>	<b>medição</b>	<b>número de amostras</b>
<b>1</b>	20236	i	6922
		ii	4967
		iii	3587
		iv	4760
<b>2</b>	20655	v	9948
		vi	10707
<b>3</b>	24944	vii	24944
<b>4</b>	29183	viii	29183

Fonte: O autor (2018)

#### 4.5 DIVISÃO DOS DADOS DE TREINAMENTO E VALIDAÇÃO

Os dados foram divididos entre dados de treinamento e validação. Neste caso, após a fase de treinamento dos modelos com diversas proporções, optou-se pelo número de amostras de treinamento por meio do cálculo da média aritmética entre três rodadas de treinamento para cada uma das técnicas empregadas e os resultados foram avaliados a partir dos dois critérios de desempenho definidos nesta dissertação, MSE e  $R^2$ . Para o modelo da rotação do ventilador as proporções entre treinamento e validação que apresentaram melhores resultados foram respectivamente 60%/40% para as técnicas SV. Para a ANN foi 70%/30% e para LSM foi 30%/70%, como apresentado na Tabela 4.5:

Tabela 4.5 – Proporção dados treinamento e validação para rotação do ventilador

Dados		SVR		LS-SVM		ANN		LSM	
Treinamento	Validação	MSE	$R^2$	MSE	$R^2$	MSE	$R^2$	MSE	$R^2$
70%	30%					X	X		
60%	40%	X	X	X	X				
40%	60%								
30%	70%							X	X

MSE refere-se ao erro dos estimadores e  $R^2$  à qualidade do modelo.

Fonte: O autor (2018)

A distribuição dos dados entre as fases de treinamento e validação para o modelo da temperatura do fluido de arrefecimento, que apresentaram melhores resultados em termos de MSE e  $R^2$ , foram 60% no treinamento e 40% para validação para LS-SVM, 70%/30% para SVR e ANN e 30%/70% para LSM, conforme ilustrado na Tabela 4.6:

Tabela 4.6 – Proporção dados treinamento e validação para temperatura do fluido de arrefecimento

Dados		SVR		LS-SVM		ANN		LSM	
Treinamento	Validação	MSE	$R^2$	MSE	$R^2$	MSE	$R^2$	MSE	$R^2$
70	30	X	X			X	X		nc
60	40			X	X				nc
40	60								nc
30	70							X	nc

MSE refere-se ao erro dos estimadores e  $R^2$  à qualidade do modelo.

nc - modelo não convergiu

Fonte: O autor (2018)

Após conhecer quais distribuições de dados para treinamento e validação permitiam o melhor desempenho dos modelos em cada técnica, optou-se em manter a relação 70%/30% em todas as técnicas, pois, satisfazia a maioria dos algoritmos de ambos modelos (temperatura e rotação) e assim poder comparar todas a técnicas, os resultados gerados são apresentados no Capítulo 5.

#### 4.6 AJUSTES ESPECÍFICOS DOS PARÂMETROS DE CADA TÉCNICA

Alguns ajustes específicos dos parâmetros de controle de cada técnica foram necessários para a realização da modelagem.

##### 4.6.1 Função *kernel* para StatLSSVM

A técnica StatLSSVM suporta cinco tipos diferentes de *kernels* definidos (semi-)positivos conforme De Brabanter et al. (2013), sendo estes: função de base radial (do inglês *Radial Basis Function – RBF\_kernel*), função linear (*lin\_kernel*), função Gaussiana (*gauss\_kernel*), Gaussiana de quarta ordem (*gauss4\_kernel*) além da Gaussiana aditiva (*gaussadd\_kernel*).

Dentre os cinco *kernels* testados para se encontrar os SV, escolheu-se a função RBF, pois, apresentou melhor desempenho entre os *kernels* em termos dos coeficientes de desempenho definidos nesta dissertação, conforme apresentado na Tabela 4.7 e para o sinal de rotação e na Tabela 4.8 para o sinal de temperatura.

Para rotação os parâmetros de ajuste da técnica foram mantidos em  $\gamma=1$  e  $h=0,3$  (regularização e largura de banda do *kernel* respectivamente), 70%/30% treinamento/validação.

Tabela 4.7 – *Kernels* para a StatLSSVM no caso da rotação ventilador

Função <i>Kernel</i>	Gaussiana	RBF	Gaussiana de 4ª ordem	Gaussiana aditiva	Linear
<b>MSE</b>	0,003	<b>0,002</b>	0,003	0,002	0,004
<b>R<sup>2</sup></b>	0,472	<b>0,766</b>	0,497	0,625	0,419
<b>tempo [s]</b>	111	88	96	328	66

Fonte: O autor (2018)

Para temperatura os parâmetros de ajuste da técnica foram mantidos em  $\gamma=1$  e  $h=0,1$  (regularização e largura de banda do *kernel* respectivamente), A relação treinamento/validação foi de 70%/30%.

Tabela 4.8 – *Kernels* para a StatLSSVM no caso da temperatura do fluido

Função <i>Kernel</i>	Gaussiana	RBF	Gaussiana de 4ª ordem	Gaussiana aditiva	Linear
<b>MSE</b>	0,026	<b>0,022</b>	0,025	0,030	0,042
<b>R<sup>2</sup></b>	0,426	<b>0,519</b>	0,443	0,349	0,090
<b>tempo [s]</b>	95,22	90,5	100,64	299,1	62,16

Fonte: O autor (2018)

#### 4.6.2 Parâmetros de ajuste para a StatLSSVM

A definição dos valores (Tabela 4.9) para os parâmetros de ajuste  $\gamma$ =regularização e  $h$ =largura de banda do *kernel* foi realizada com a função “*tunelssvm*” da ferramenta LSSVM para uma relação treinamento/validação de 60%/40% e *kernel* RBF, como definido anteriormente nos Capítulos 4.5 e 4.6.1.

Tabela 4.9 – Parâmetros de ajuste para a StatLSSVM

Variável/Modelo	$\gamma$	$h$	tempo de processamento [h]
<b>rotação</b>	72,1155	0,094	7,0
<b>temperatura</b>	6,1059	0,021	7,5

Fonte: O autor (2018)

#### 4.6.3 Função *kernel* para SVR

Para a SVR as funções *kernel* disponíveis foram RBF, gaussiana, polinomial. A que apresentou melhor desempenho foi a SVR com RBF, conforme apresentado na Tabela 4.10 para rotação e Tabela 4.11 para temperatura, com uma partição de 70% dos dados para treinamento e 30% para a validação. Com relação ao modelo SVR empregando uma *kernel* polinomial, o modelo não convergiu de forma que os valores obtidos de  $R^2$  e MSE foram incoerentes (MSE = 11.57 e  $R^2 = -1795.90$ ). Para  $R^2$  valores mais próximos de 1 indicarão um modelo próximo à perfeição ao passo que quanto menor o valor apresentado por MSE indicará o menor erro alcançado para os estimadores dos modelos.

Tabela 4.10 – Função *kernel* para a velocidade do ventilador

<b>Critério</b>	<b>RBF</b>	<b>Gaussiano</b>
<b>MSE</b>	<b>0,0024</b>	0,003
<b><math>R^2</math></b>	<b>0,6307</b>	0,5875
<b>tempo [s]</b>	<b>16,99</b>	16,93

Fonte: O autor (2018)

Tabela 4.11 – Função *kernel* para a temperatura do fluido

<b>Critério</b>	<b>RBF</b>	<b>Gaussiano</b>
<b>MSE</b>	<b>0,0261</b>	0,0262
<b><math>R^2</math></b>	<b>0,4309</b>	0,4290
<b>tempo [s]</b>	<b>6,25</b>	6,34

Fonte: O autor (2018)

#### 4.6.4 Função de ativação para ANN

Para ANN a função de ativação escolhida foi a Bayesiana, dentre as funções *Levenberg-Marquardt*, *Scaled Conjugate Gradient* e *Residual Backpropagation*. O melhor desempenho para rotação do ventilador foi utilizando-se a função Bayesiana com 20 neurônios conforme apresentado na Tabela 4.12.

Tabela 4.12 – Funções de ativação para ANN – rotação do ventilador (20 neurônios)

<b>Critério</b>	<b>Levenber-Marquardt</b>	<b>Bayesiana</b>	<b>Scaled Conjugated Gradient</b>	<b>Resilient Backpropagation</b>
<b>MSE</b>	0,0015	<b>0,0013</b>	0,0025	0,0022
<b><math>R^2</math></b>	0,7608	<b>0,8076</b>	0,6196	0,6627
<b>tempo [s]</b>	4,34	49,93	2,70	3,20

Fonte: O autor (2018)

O melhor desempenho tanto em termos de MSE quanto  $R^2$  para temperatura do fluido foi utilizando-se a função Bayesiana com 50 neurônios conforme apresentado na Tabela 4.13.

Tabela 4.13 – Funções de ativação para ANN – temperatura do fluido (50 neurônios)

Critério	Levenber-Marquardt	Bayesiana	Scaled Conjugated Gradient	Resilient Backpropagation
<b>MSE</b>	0,0894	<b>0,0223</b>	0,1179	0,0275
<b>R<sup>2</sup></b>	0,5025	<b>0,5130</b>	0,3426	0,4025
<b>tempo [s]</b>	16,17	178,33	4,60	6,18

Fonte: O autor (2018)

#### 4.6.5 Número de neurônios para ANN

A rede neural perceptron multicamadas foi treinada com 10, 20 e 50 neurônios e os resultados indicaram que o melhor desempenho de treinamento foi com 20 neurônios para rotação apesar de o  $R^2$  para rotação com 50 neurônios ter sido ainda melhor o seu tempo de processamento foi quase quatro vezes maior do que com 20 neurônios, conforme apresentado na Tabela 4.14.

Tabela 4.14 – Número de neurônios para ANN – rotação do ventilador

Critério	10	20	50
<b>MSE</b>	0,0016	<b>0,0013</b>	0,0013
<b>R<sup>2</sup></b>	0,7614	<b>0,8076</b>	0,8206
<b>tempo [s]</b>	16,41	<b>49,93</b>	189,07

trein/val: 70/30% | Função ativação: Bayesiana

Fonte: O autor (2018)

No caso da temperatura o melhor desempenho foi com e 50 neurônios conforme apresentado na Tabela 4.15:

Tabela 4.15 – Número de neurônios para ANN – temperatura do fluido

Critério	10	20	50
<b>MSE</b>	0,0258	0,0247	<b>0,0223</b>
<b>R<sup>2</sup></b>	0,4377	0,4635	<b>0,5130</b>
<b>tempo [s]</b>	14,40	29,67	178,33

trein/val: 70/30% | Função ativação: Bayesiana

Fonte: O autor (2018)

## 5 RESULTADOS

Para efeito de comparação adotou-se, em todas as técnicas a distribuição de 70% e 30% para os dados de teste e validação respectivamente. Sendo assim, o trajeto utilizado foi o número 1 com 20236 amostras, conforme definido no Capítulo 4.4, sendo 14165 para treinamento dos modelos e 6070 para validação. As saídas dos modelos foram deslocadas em  $(t-1)$  com relação as suas entradas.

### 5.1 TESTE DE ESTABILIDADE DOS MODELOS

Após definidos os parâmetros específicos para cada técnica, bem como a melhor distribuição dos dados para treinamento e validação, como apresentado no Capítulo 4.5, foram rodadas três sequencias de treinamento e validação para cada uma das técnicas conforme apresentado nas Tabelas 5.1 e 5.2, para rotação e temperatura respectivamente.

Tabela 5.1 – Treinamento e validação das técnicas para a rotação do ventilador

Critério	LSM			SVR			StatLSSVM			ANN		
MSE	0,0036	0,0040	0,0037	0,0026	0,0022	0,0027	0,0011	0,0010	0,0009	0,0014	0,0013	0,0012
R2	0,464	0,428	0,450	0,612	0,644	0,624	0,840	0,840	0,860	0,805	0,788	0,830
tempo[s]	0,4	0,5	0,3	26,3	19,1	19,7	40,1	39,3	39,3	60,4	38,0	51,4
Treinamento/Validação: 70/30%										20 neurônios		

Fonte: O autor (2018)

Tabela 5.2 – Treinamento e validação das técnicas para a temperatura do fluido

Critério	LSM			SVR			StatLSSVM			ANN		
MSE	0,0073	0,0073	0,0073	0,0261	0,0629	0,0276	0,0176	0,0175	0,0173	0,0223	0,0226	0,0220
R2	-0,586	-0,572	-0,569	0,431	0,427	0,416	0,606	0,615	0,623	0,513	0,498	0,528
tempo[s]	0,3	0,3	1,0	22,3	20,5	21,1	49,1	46,3	48,2	184,0	176,0	175,0
Treinamento/Validação: 70/30%										50 neurônios		

Fonte: O autor (2018)

Os resultados apresentados nas tabelas abaixo referem-se à média aritmética de cada um dos critérios de desempenho em cada técnica, sendo a Tabela 5.3 a média aritmética dos resultados obtidos da Tabela 5.1, relativos aos modelos da rotação do ventilador.

Tabela 5.3 – Comparativos da validação para rotação do ventilador

<b>Critério</b>	<b>LSM</b>	<b>SVR</b>	<b>StatLSSVM</b>	<b>ANN</b>
<b>MSE</b>	0,0040	0,0025	<b>0,0011</b>	0,0013
<b><math>R^2</math></b>	0,4348	0,6266	<b>0,8357</b>	0,8076
<b>tempo [s]</b>	<b>0,4</b>	22	40	50

Fonte: O autor (2018)

Da mesma forma, para a obtenção dos resultados para a temperatura do fluido de arrefecimento, na Tabela 5.4 tem-se a média aritmética dos resultados obtidos a partir da Tabela 5.2.

Tabela 5.4 – Comparativos da validação para temperatura do fluido

<b>Critério</b>	<b>LSM</b>	<b>SVR</b>	<b>StatLSSVM</b>	<b>ANN</b>
<b>MSE</b>	0,0758	0,0389	<b>0,0179</b>	0,0223
<b><math>R^2</math></b>	-0,8604	0,4242	<b>0,6115</b>	0,5130
<b>tempo [s]</b>	0,5	<b>21</b>	48	178

Fonte: O autor (2018)

Pôde-se observar que as técnicas de Aprendizado de máquina sobressaíram-se em relação a técnica clássica de modelagem LSM, pelos respectivos desempenhos calculados pelo coeficiente  $R^2$  que determina a qualidade do modelo. O resultado da LSM foi inferior ao das técnicas de AM para os modelos de rotação e não convergiu para os modelos de temperatura.

Com relação a técnicas de AM, pode-se observar que a StatLSSVM resultou em melhores desempenhos dentre os modelos de rotação e temperatura. Para a rotação os resultados foram de  $R^2=0,8357$  contra  $R^2=0,8076$  para ANN (Tabela 5.3). Com relação ao erro, que garante a qualidade dos estimadores, a StatLSSVM obteve o menor MSE=0,0011 seguido pelas ANN que apresentaram um erro MSE=0,0013. O tempo de processamento para StatLSSVM foi 20% menor

do que o tempo de processamento demandado pelas ANN, que demandaram 50 segundos médios de processamento como apresentado na Tabela 5.3.

Para os modelos de temperatura a StatLSSVM também obteve os melhores resultados. Em termos de desempenho,  $R^2=0,6115$  e de erro  $MSE=0,0179$ , com um tempo de processamento de 48 segundos, 27% do tempo gasto pelas ANN, cujo segundo melhor resultado entre as técnicas foi de  $R^2=0,5130$  e  $MSE=0,0223$  como apresentado na Tabela 5.4.

## 5.2 TREINAMENTO E VALIDAÇÃO

A seguir são apresentados os resultados para treinamento e validação em cada técnica, sendo mantida a proporção 70/30% respectivamente para todas as técnicas.

A técnica LSM foi a mais rápida com um tempo processamento de 0,5 segundo para treinamento e validação do modelo de rotação do ventilador apresentando um  $R^2$  de 0,4509 para validação do modelo. A técnica StatLSSVM por sua vez obteve o melhor desempenho na validação com  $R^2$  de 0,8410 e um tempo de 37 segundos, as ANN obtiveram um  $R^2$  de 0,7907 com 19 segundos de processamento. Finalizando, a SVR apresentou um  $R^2$  igual a 0,6238 de validação do modelo e tempo de 13 segundos, conforme apresentado na Tabela 5.5.

Tabela 5.5 – Comparativos de treinamento e validação para rotação do ventilador

Critério	LSM		SVR		StatLSSVM		ANN	
	treinamento	validação	treinamento	validação	treinamento	validação	treinamento	validação
<b>MSE</b>	0,0039	0,0039	0,0025	0,0027	0,0008	0,0011	0,0012	0,0013
<b>R<sup>2</sup></b>	0,4099	0,4509	0,6251	0,6238	0,8782	0,8410	0,8288	0,7907
<b>tempo [s]</b>	0,5		13		37		19	

Fonte: O autor (2018)

Para a temperatura do fluido, a técnica SVR foi a mais rápida com um tempo processamento de 10 segundos para treinamento e validação do modelo, porém apresentando um  $R^2$  de validação igual a 0,4290. A técnica StatLSSVM, por sua vez, obteve o melhor desempenho na validação do modelo com um  $R^2$  de

0,6182 e tempo de processamento de 50 segundos, as ANN obtiveram um  $R^2$  de 0,4739 na validação com 92 segundos. O modelo de temperatura não convergiu para LSM apresentando valores incoerentes para  $R^2$ , conforme a Tabela 5.6.

Tabela 5.6 – Comparativos de treinamento e validação para temperatura do fluido

Critério	LSM		SVR		StatLSSVM		ANN	
	treinamento	validação	treinamento	validação	treinamento	validação	treinamento	validação
MSE	0,0732	0,0717	0,0254	0,0264	0,0110	0,0171	0,0203	0,0244
$R^2$	-0,5814	-0,5831	0,4463	0,4290	0,7700	0,6182	0,5573	0,4739
tempo [s]	0,6		10		50		92	

Fonte: O autor (2018)

Resumindo, a StatLSSVM foi a técnica com o melhor desempenho  $R^2$  nos dois modelos, sendo 84% para rotação do ventilador e 62% para o modelo de temperatura.

Na sequência,  $R^2$  para rotação, ficaram ANN com 79%, SVR com 62% e LSM 45%. Nota-se a mesma sequência com relação ao MSE: StatLSSVM igual a 0,0011, ANN com 0,0013, SRV com 0,0027 e LSM com 0,0039.

A mesma tendência observou-se para o modelo da temperatura do fluido, sendo que a melhor técnica foi a StatLSSVM com  $R^2$  de 62%, ANN com 47% e SVR com 43%, LSM não convergiu apresentando resultados inconsistentes para  $R^2$ .

Com relação ao MSE para a temperatura do fluido, a StatLSSVM apresentou o menor erro de validação igual a 0,0171, seguido pela ANN com 0,0244 e SVR com 0,0264.

### 5.3 REPRESENTAÇÃO GRÁFICA DOS RESULTADOS

Para todos os resultados apresentados seja a partir da média aritmética de três medições com cada técnica em ambos os modelos ou mesmo nos testes de treinamento e validação apresentados neste capítulo, pôde-se observar que os algoritmos StatLSSVM apresentaram os melhores resultados de qualidade dos modelos, medido através de  $R^2$  retornando o menor erro na qualidade dos estimadores apresentados por meio do MSE.

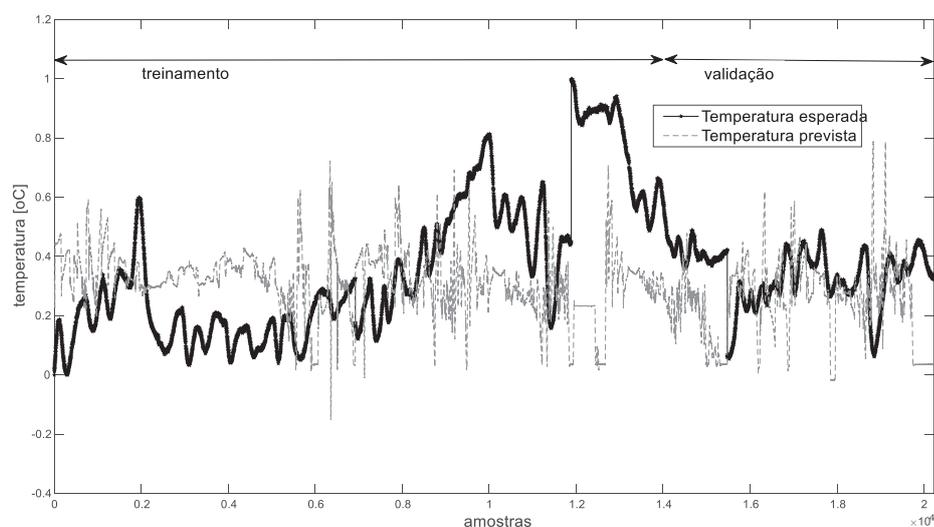
Nas Figuras 5.1 a 5.16 são apresentados os melhores resultados de modelagem de cada uma das técnicas, ou seja, LSM, StatLSSVM, SVR e ANN respectivamente.

### 5.3.1 Modelagem LSM

O método dos mínimos quadrados mostrou-se útil para se trabalhar inicialmente os dados de entrada escolhidos e verificar a repetibilidade dos resultados obtidos nos diferentes trajetos utilizados, descritos no Capítulo 4. Embora robusto e com um erro MSE relativamente pequeno o LSM não se mostrou eficiente na qualidade do modelo gerado indicado pelo índice  $R^2$  que ficou em torno de 40% para o sinal de rotação (Figura 5.1) e não convergiu para o modelo da temperatura, como apresentado na Figura 5.3.

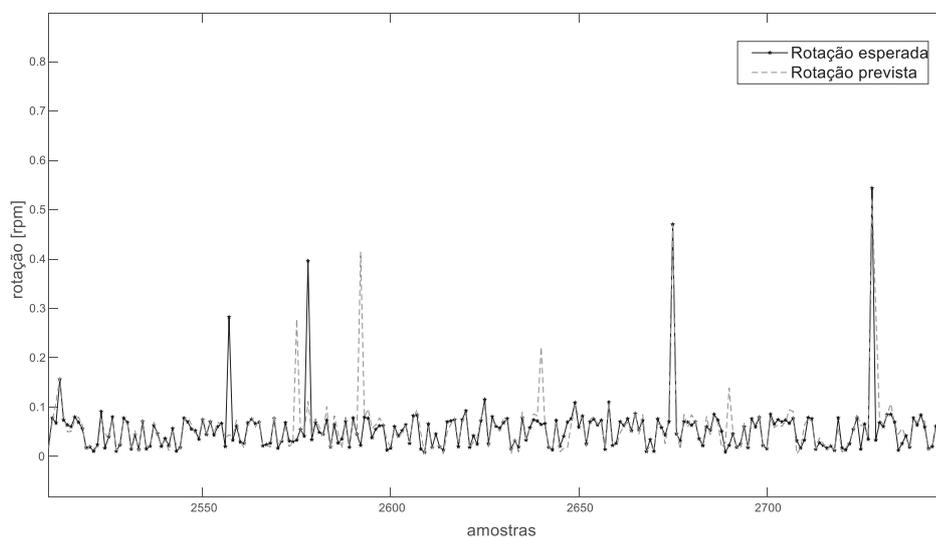
Segundo o Capítulo 3.2.1, o LSM possui convergência lenta e depende da dimensionalidade dos dados para estabilizar, uma quantidade maior de dados adquiridos no trajeto desenvolvido pelo veículo de referência poderia contribuir para a melhora do desempenho do LSM, no entanto, para os objetivos desta dissertação os resultados apresentados foram suficientes na motivação pelo uso das técnicas propostas de AM.

Figura 5.1 - Rotação do ventilador com LSM



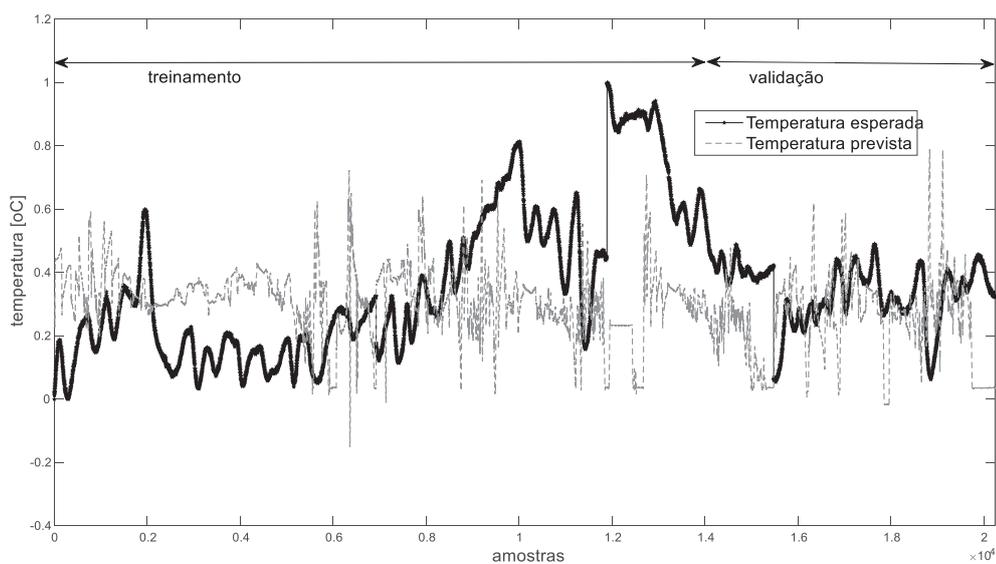
Os dados brutos foram distribuídos aleatoriamente, conforme informado no Capítulo 4.4 e apresentados na Figura 5.2 (trecho destacado aleatoriamente).

Figura 5.2 – Validação rotação do ventilador para LSM



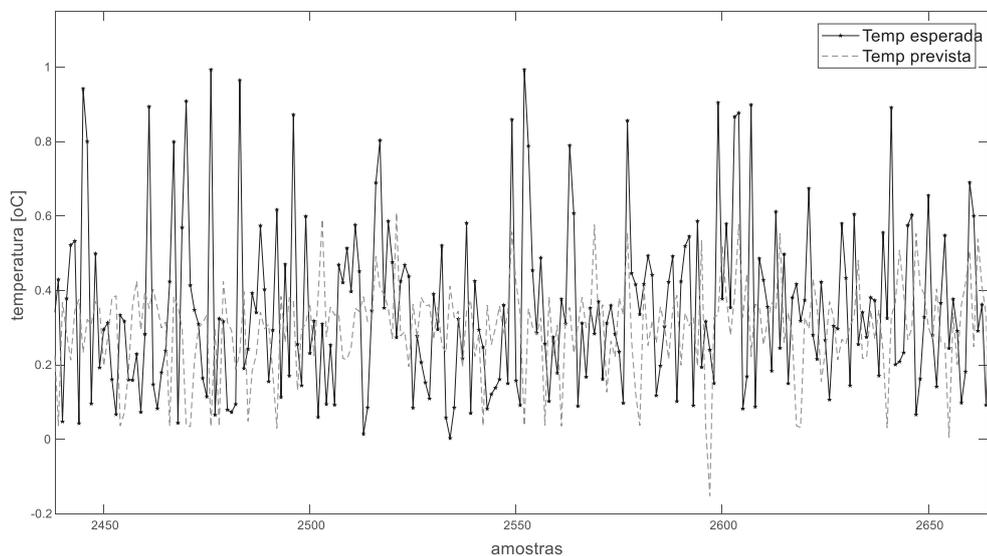
Na Figura 5.3 é apresentado o modelo da temperatura. Nota-se que o mesmo não convergiu, como apresentado na Tabela 5.4

Figura 5.3 – Temperatura do fluido com LSM



Os dados brutos foram distribuídos aleatoriamente, conforme informado no Capítulo 4.4 e apresentados na Figura 5.4 (trecho destacado aleatoriamente).

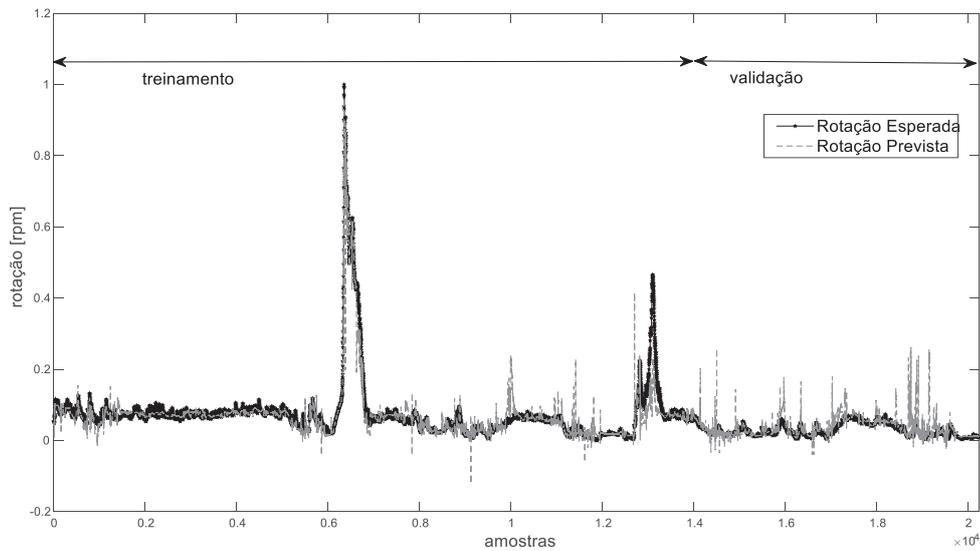
Figura 5.4 – Validação temperatura do fluido para LSM



### 5.3.2 Modelagem StatLSSVM

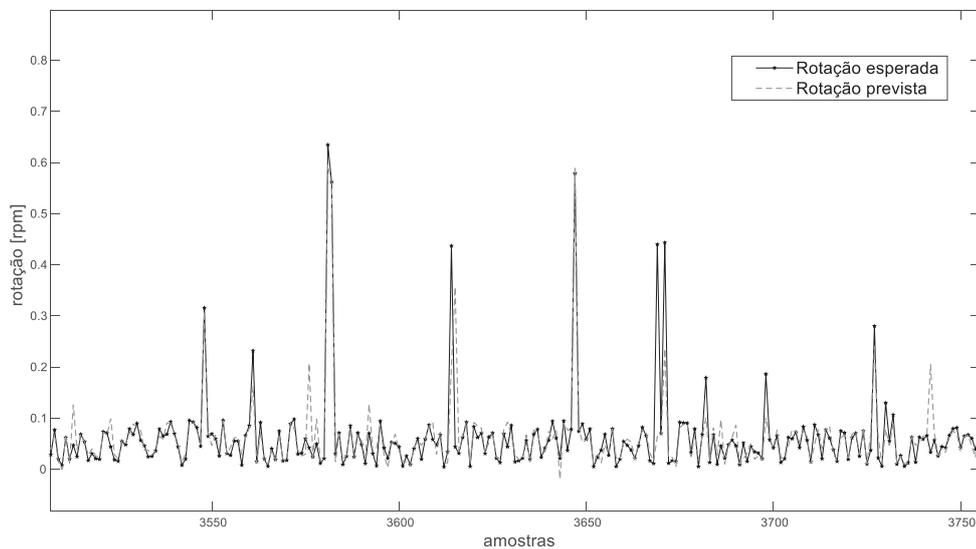
As técnicas de vetores de suporte se mostraram de fato eficientes entre as investigadas neste trabalho mas quando utilizada com mínimos quadrados se mostrou a mais eficiente como observado na Figura 5.5 com um  $R^2$  de 87% no treinamento e na validação 84% (Tabela 5.5), o melhor de todos os resultados, estes otimizados pela utilização da função *kernel* de base radial definida como a melhor opção para os modelos SVM desta dissertação e parâmetros  $\gamma$ =regularização e  $h$ =largura de banda obtidos pela função *tuning* disponibilizado pela StatLSSVM, detalhes em 4.6.1 e 4.6.2 respectivamente.

Figura 5.5 – Rotação do ventilador com StatLSSVM



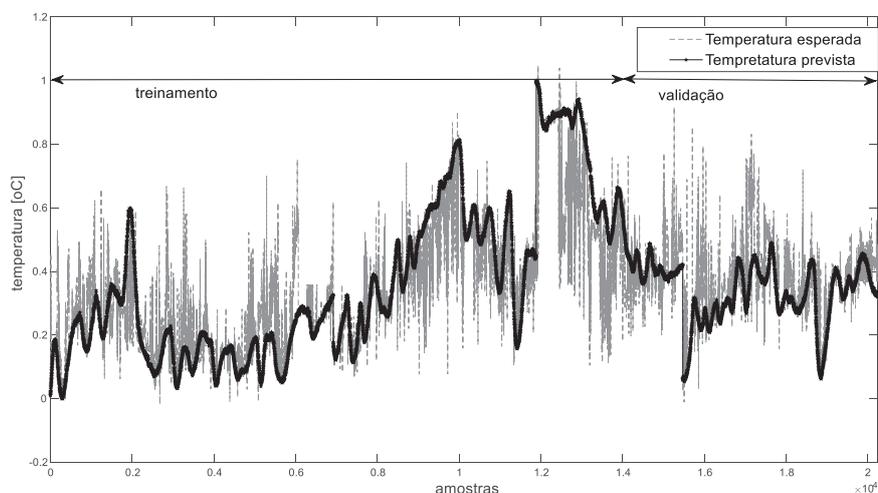
Os dados brutos foram distribuídos aleatoriamente, conforme informado no Capítulo 4.4 e apresentados na Figura 5.6 (trecho destacado aleatoriamente).

Figura 5.6 – Validação rotação do ventilador para StatLSSVM



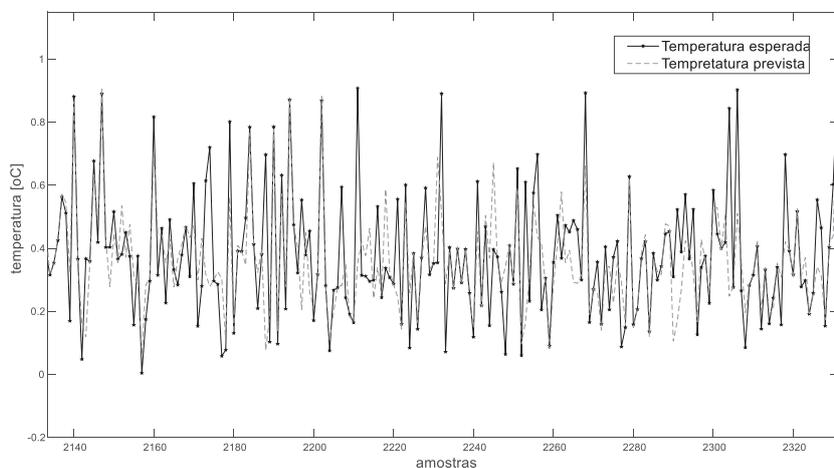
Assim como nos resultados obtidos da StatLSSVM para rotação os resultados para o modelo da temperatura do fluido de arrefecimento se mostraram os melhores entre todas as técnicas abordadas (Figura 5.7), mas ainda existe possibilidades para melhora,  $R^2$  de 77% no treinamento e na validação 62% (Tabela 5.6).

Figura 5.7 – Temperatura do fluido com StatLSSVM



Os dados brutos foram distribuídos aleatoriamente, conforme informado no Capítulo 4.4 e apresentados na Figura 5.8 (trecho destacado aleatoriamente).

Figura 5.8 – Validação temperatura do fluido para StatLSSVM

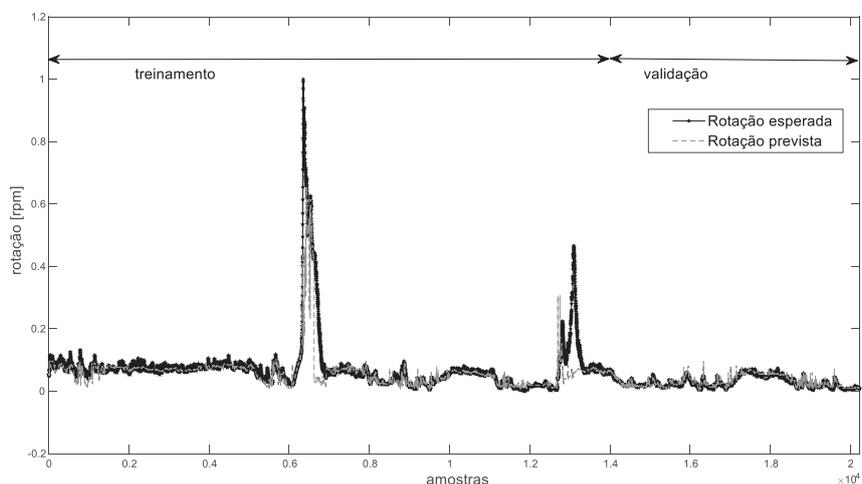


### 5.3.3 Algoritmos SVR

A SVM para regressão usando uma função *kernel* RBF não conseguiu os resultados esperados para se tornar uma ferramenta de uso no trabalho de calibração dos caminhões, quando comparadas com a StatLSSVM e também com a ANN. Os resultados para o modelo de rotação foram os mesmos tanto nas fases

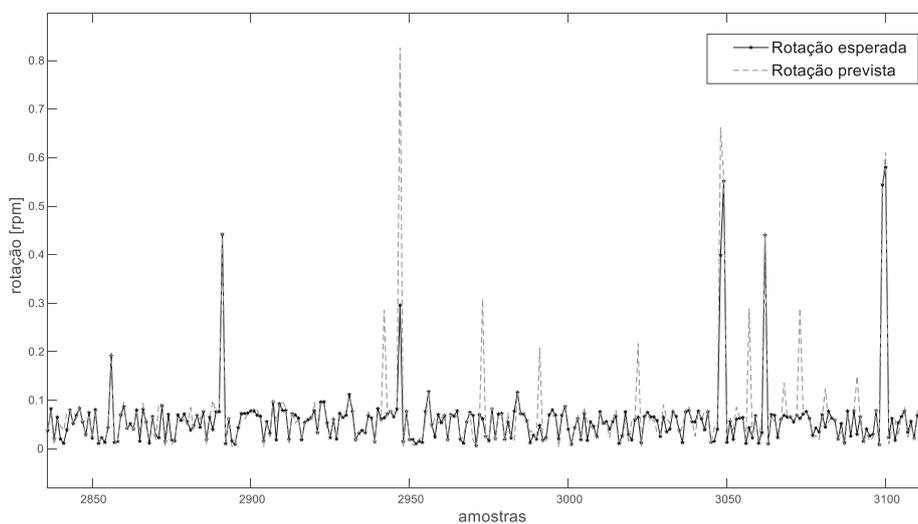
de treinamento quanto validação,  $R^2$  de 62% (Tabela 5.5), como observado na Figura 5.9.

Figura 5.9 – Rotação do ventilador com SVR



Os dados brutos foram distribuídos aleatoriamente, conforme informado no Capítulo 4.4 e apresentados na Figura 5.10 (trecho destacado aleatoriamente).

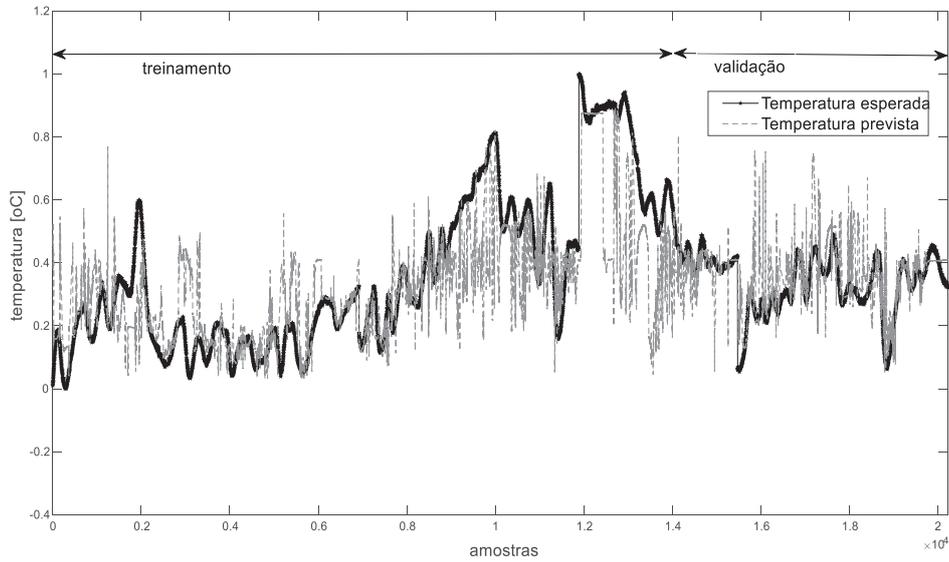
Figura 5.10 – Validação rotação do ventilador para SVR



Na Figura 5.11 a saída do modelo SVR da temperatura de fluido apresentado para um  $R^2$  de 44% no treinamento e 42% na validação. Os resultados

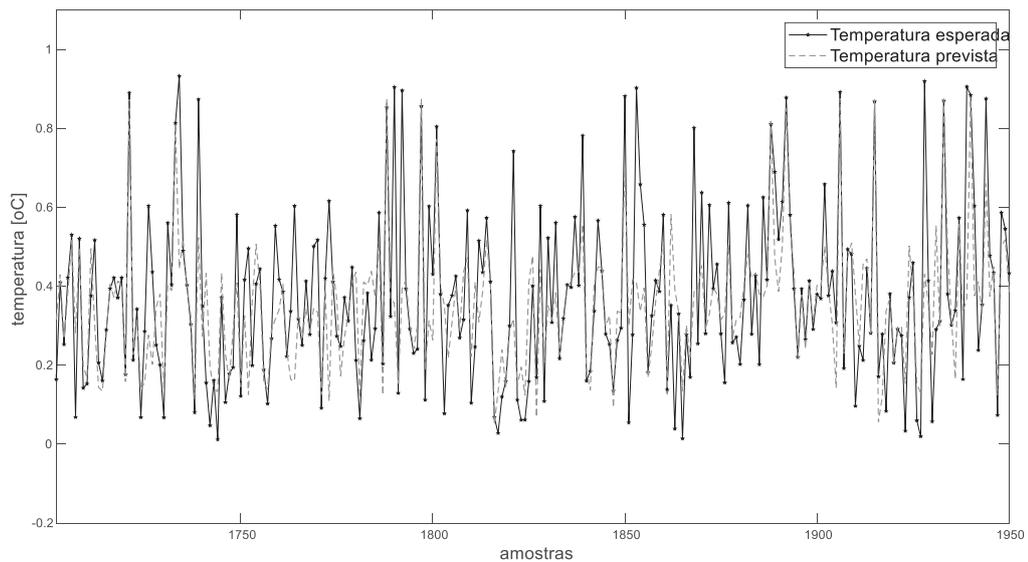
estão abaixo do esperado, apesar do tempo de processamento ser inferior entre todas as técnicas abordadas.

Figura 5.11 – Temperatura do fluido com SVR



Os dados brutos foram distribuídos aleatoriamente, conforme informado no Capítulo 4.4 e apresentados na Figura 5.12 (trecho destacado).

Figura 5.12 – Validação temperatura do fluido para SVR

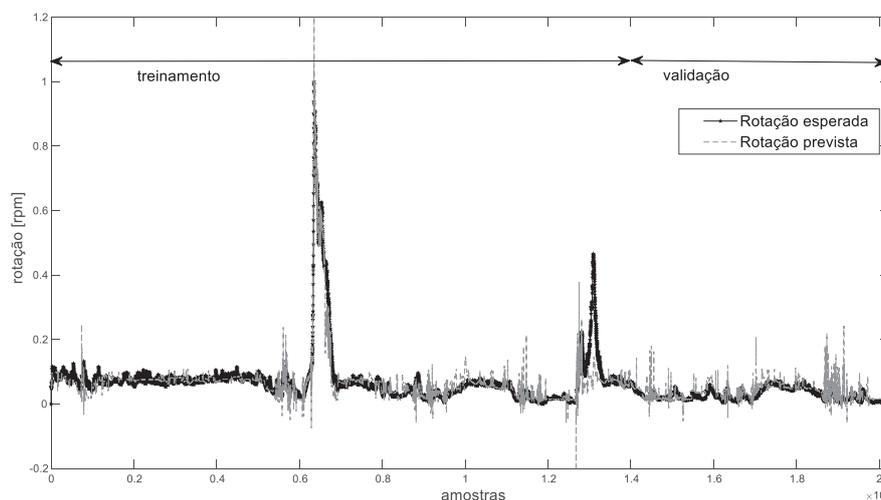


### 5.3.4 Algoritmos ANN

Os resultados alcançados pela ANN foram próximos aos da StatLSSVM, mostrando que para os modelos gerados neste trabalho, esta técnica pode ser também uma opção a ser utilizada, no entanto o tempo de processamento pode caracterizar uma restrição, principalmente para o modelo de temperatura (Tabela 5.4).

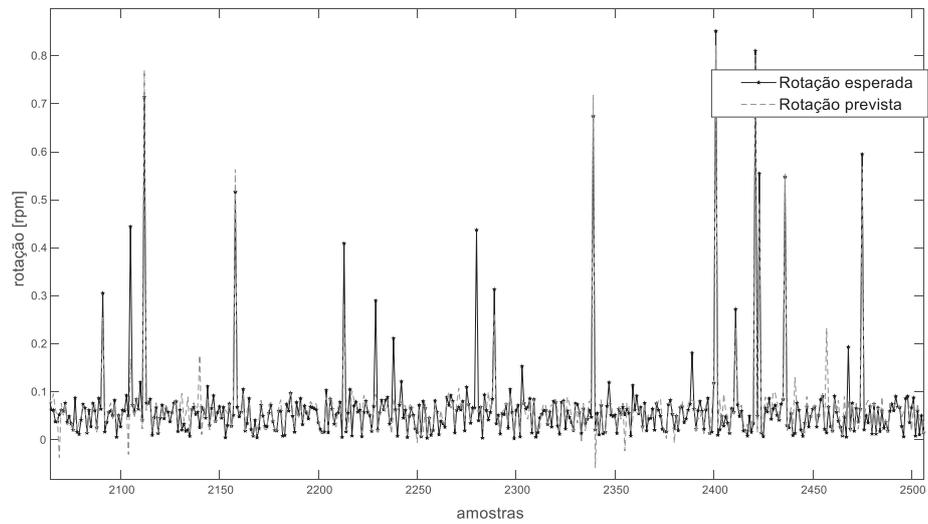
Abaixo a representação gráfica resultante da ANN para o modelo de rotação do ventilador, com uma rede de 20 neurônios e função de ativação RBF, resultando em um  $R^2$  de 83% no treinamento e 79% na validação (Tabela 5.5), conforme Figura 5.13 .

Figura 5.13 – Rotação do ventilador com ANN



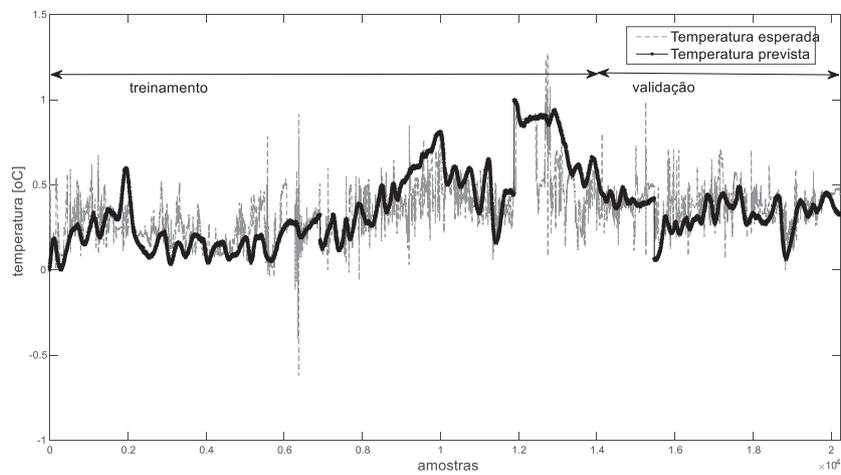
Os dados brutos foram distribuídos aleatoriamente, conforme informado no Capítulo 4.4 e apresentados na Figura 5.14 (trecho destacado aleatoriamente).

Figura 5.14 – Validação rotação do ventilador para



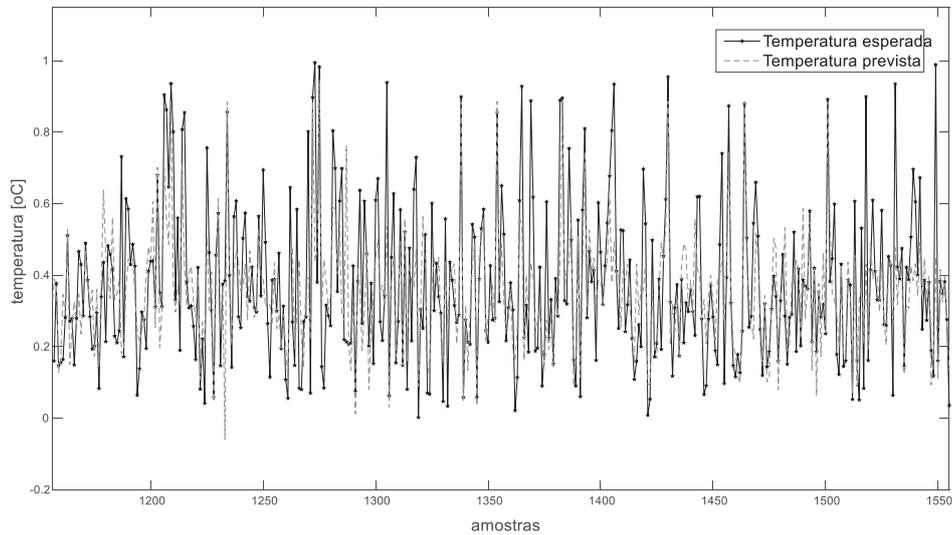
O resultado da ANN apresentado na Figura 5.15, para a modelagem da temperatura do fluido, com uma rede de 50 neurônios e função de ativação RBF, resultando em um  $R^2$  de 55% no treinamento e 47% na validação (Tabela 5.5).

Figura 5.15 – Temperatura do fluido com ANN



Os dados brutos foram distribuídos aleatoriamente, conforme informado no Capítulo 4.4 e apresentados na Figura 5.16 (trecho destacado aleatoriamente).

Figura 5.16 – Validação temperatura do fluido para ANN



## 6 CONCLUSÃO

O sistema de arrefecimento dos motores Diesel, que é um componente fundamental dos veículos pesados, tem um papel importante para garantir a operação do motor e também contribuir para a redução do consumo de combustível e de poluentes na atmosfera.

Nesta dissertação buscou-se modelar dois sinais importantes para o desempenho e eficácia da funcionalidade desse sistema: a rotação do ventilador e a temperatura do fluido de arrefecimento, cuja finalidade é fornecer entrada e atuar de maneira positiva no controle da temperatura do motor em tempo hábil para garantir tal equilíbrio. Optou-se pelo uso de técnicas de Aprendizado de máquina que conseguissem lidar com as não linearidades do sistema tal como ANN do tipo perceptron multicamadas, a SVM e um modelo linear com modelagem por meio do LSM.

Este trabalho mostrou que é possível utilizar modelos de regressão na obtenção de modelos para o sistema de arrefecimento do motor. O sistema de arrefecimento possui características complexas de se modelar, a escolha dos sinais de entrada foi suficiente para obter os resultados dos modelos e validar as técnicas escolhidas, a possibilidade de se testar diferentes medições, adquiridas em trajetos reais realizados pelo veículo de referência fizeram com que os resultados pudessem ser representativos para o comportamento do sistema de arrefecimento e reproduzido pelos modelos de temperatura e rotação gerados em cada uma das técnicas utilizadas, enfim, todos os objetivos propostos foram atendidos com sucesso.

Os resultados obtidos da modelagem do sistema de arrefecimento foram com  $R^2$  acima de 60% para o sinal de temperatura e acima de 80% para o sinal de rotação, indicando que existe espaço para melhorias. O exercício da correlação dos sinais de entrada do modelo com suas saídas confirmou que, principalmente no caso da temperatura do fluido do motor, é possível obter melhor desempenho pela redução dos sinais de entrada e conseqüentemente a simplificação do modelo (Capítulo 4.2). O estudo do sistema indicou que outros sinais poderiam ser inseridos ao modelo como a potência consumida pelo ventilador em diferentes rotações de trabalho e o deslizamento do fluido nos momentos de acoplamento do ventilador.

A utilização de outros trajetos e uma maior quantidade de dados de campo, também poderiam contribuir para a melhora do desempenho dos modelos.

A aplicação dos modelos na prática é apenas o início de um trabalho maior no desenvolvimento de futuros algoritmos para o sistema de arrefecimento os quais poderiam ser integrados ao software automotivo. Como resultado final desta dissertação é esperado uma facilitação no trabalho dos engenheiros de calibração pelo conhecimento prévio do comportamento do sistema nas diferentes aplicações a que os caminhões deverão operar, sendo de grande valor em termos de tempo gasto com expedições a diferentes ambientes e centenas de horas de tentativas de calibração em busca do melhor comportamento do sistema em diferentes situações. Uma vez definida a técnica mais apropriada e de posse dos dados de aplicação de um certo veículo, bastaria apenas rodar os modelos e produzir novas calibrações a partir dos resultados previstos, reduzindo o tempo gasto para a refinação do sistema. Neste caso a ferramenta StatLSSVM poderia ser empregada, pois, é de fácil utilização e não demanda conhecimentos avançados de AM, uma vez entendido seus comandos e ajustes iniciais.

A partir deste trabalho, como mencionado anteriormente, um estudo aprofundado de condições de contorno do sistema e a utilização de técnicas de otimização capazes de refinar os dados poderiam contribuir na geração de modelos robustos, conseqüentemente mais realistas, sobre o funcionamento do sistema de arrefecimento. Algoritmos para a previsão de suas reações em tempo real poderiam ser também desenvolvidos e assim ser incorporados à lógica do fabricante do motor.

Como trabalhos futuros são sugeridos a consolidação da capacidade do modelo, tanto em prever não apenas a rotação do ventilador ou a temperatura, mas a previsão de outras informações que beneficiariam os trabalhos de calibração do sistema e a futura integração destes modelos ao software embarcado da ECU do motor.

## REFERÊNCIAS

- AGUIRRE, L. A. **Introdução à identificação de sistemas. Técnicas lineares e não lineares: Teoria e aplicação**. 4a ed. Belo Horizonte, Brasil: Editora Ufmg, 2015.
- AGUIRRE, M. V.; CORRÊA, L. Identificação não-linear caixa-cinza: uma revisão e novos resultados. **Sba: Controle & Automação Sociedade Brasileira de Automatica**, v. 15, nr. 2, n. ISSN 0103-1759, 2004.
- ASHOK, B.; ASHOK, S. D.; KUMAR, C. R. A review on control system architecture of a SI engine management system. **Annual Reviews in Control**, v. 41, p. 94–118, 2016. Elsevier Ltd.
- BEHR, H. S. G. Sophisticated: Engine Cooling for Commercial Vehicles. **Behr Hella Service**, 2013. Germany.
- BOSCH, R. G. **Manual de Tecnologia Automotiva BOSCH**. 25a ed. São Paulo, Brasil: Editora Edgard Blücher, 2004.
- BOSE, B. E.; GUYON, I. M.; VAPNIK, V. N. A Training Algorithm for Optimal Margin Classifiers. **Proceedings of the fifth annual workshop on Computational learning theory**, p. 144–152, 1992.
- DE BRABANTER, K.; SUYKENS, J. A. K.; DE MOOR, B. Nonparametric regression via StatLSSVM. **Journal of Statistical Software**, v. 55, n. 2, p. 1–21, 2013.
- BURGES, C. A Tutorial on Support Vector Machines for Pattern Recognition. **Data Min. Knowl. Discov.**, v. 2, n. 2, p. 121–167, 1998.
- CAMPOS, E. R. V. **Modelagem do uso de nanofluidos no sistema de arrefecimento de motores a combustão interna**, 2010. Pontifícia Universidade Católica do Rio de Janeiro - PUC-RIO.
- CONAMA IBAMA. **Diesel e emissões - A nova Legislação 2012**. São Paulo, Brasil: ANFAVEA, 2017.
- CORTES, C.; VAPNIK, V. Support-Vector Networks. **Machine Learning**, v. 20, p. 273–297, 1995.
- GONÇALVES, A. R. **Máquina de Vetores Suporte**. , 2009. Campinas, Brasil: Unicamp.DCA.FEE.
- GRELET, V.; DUFOUR, P.; NADRI, M.; et al. Model based control for waste heat recovery heat exchangers Rankine cycle system in heavy duty trucks. **Proceedings of the 3rd International Seminar on ORC Power Systems. Brussels, Belgium**, v. 191, n. 2007, p. 1–10, 2015.
- HAGHIGHAT, A. K., ROUMI, S., MADANI, N., BAHMANPOUR, D., & OLSEN, M. G. An intelligent cooling system and control model for improved engine thermal management. **Applied Thermal Engineering**, 128, 253–263, 2018.
- HAYKIN, S. **Redes Neurais - Princípios e Prática**. 2a ed. Ontario, Canadá:

Bookman, 2001.

HEYWOOD, J. B. **Internal combustion engines fundamentals**. New York, USA: McGraw Hill. Inc., 1988.

HORTA, E. G. **Aplicação de Máquinas de Aprendizado Extremo ao Problema de Aprendizado Ativo**, 2015. Universidade Federal de Minas Gerais - Belo Horizonte.

HUANG, T.-M.; KECCMAN, V.; KOPRIVA, I. **Kernel Based Algorithms for Mining Huge Data Sets: Supervised, Semi-supervised, and Unsupervised Learning (Studies in Computational Intelligence)**. 1st ed. Warsaw, Poland: Springer, 2010.

INCOSE, I. C. ON S. E. **Systems engineering handbook: a guide for system life cycle processes and activities**. v. 3.2.2 ed. San Diego. USA: INCOSE, 2011.

ISERMANN, R., & HAFNER, M. Mechatronic Combustion Engines - from Modeling to Optimal Control. **European Journal of Control**, (7), 220–247, 2001.

ISKANDAR, M. A.; ADADE FILHO, A. Design and analysis of a cooling system of a diesel engine, emissions and fuel consumption. **ABCM Symposium Series in Mechatronics**, v. 5, p. 39–48, 2013.

KARIMPOUR, M.; HITIHAMILLAGE, L.; ELKHOORY, N.; MORIDPOUR, S.; HESAMI, R. Fuzzy Approach in Rail Track Degradation Prediction. **Journal of Advanced Transportation**, v. 2018, 2018.

KHODABAKHSHIAN, M.; FENG, L.; WIKANDER, J. Predictive control of the engine cooling system for fuel efficiency improvement. **2014 IEEE International Conference on Automation Science and Engineering (CASE)**, p. 61–66, 2014.

KUNG, S. Y. **Kernel Methods and Machine Learning**. Princeton, USA: Cambridge University Press, 2014.

LIMA, C. A. D. M. **Comitê de Máquinas: Uma Abordagem Unificada Empregando Máquinas de Vetores-Suporte**, 2004. Universidade Estadual de Campinas, UNICAMP.

LJUNG, L. **System Identification Theory for the User**. 2nd ed. Linköping, Sweden: Prentice Hall PTR, 1999.

LJUNG, L.; SÖDERSTRÖM, T. **Theory and Practice of Recursive Identification**. Cambridge, USA: MIT Press, 1983.

LORENA, ANA C. E CARVALHO, A. C. P. L. F. **Introdução aos Classificadores de Margens Largas**. São Carlos, SP, Brasil, 2003.

LORENA, A. C.; CARVALHO, A. C. P. L. F. DE. Uma Introdução às Support Vector Machines. **Revista de Informática Teórica e Aplicada, RITA**, v. 14, n. 2, p. 43–67, 2007.

MERCER, J. Functions of Positive and Negative Type, and their Connection with the Theory of Integral Equations. **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, v. 209, n. 441–

458, p. 415–446, 1909.

MERUELO, A. C.; SIMPSON, D. M.; VERES, S. M.; NEWLAND, P. L. Improved system identification using artificial neural networks and analysis of individual differences in responses of an identified neuron. **Neural Networks**, v. 75, p. 56–65, 2016. Elsevier Ltd.

NELLES, O. **Nonlinear system identification: from classical approaches to neural networks and fuzzy models**. Berlin, Germany: Springer-Verlag New York Bertin Heidelberg, 2001.

OGUNFUNMI, T. **Adaptive Nonlinear System Identification, The Volterra and Wiener Model Approaches**. Santa Clara, CA, USA: Springer Science+Business Media, LLC, 2007.

PANG, H. H.; BRACE, C. J.; AKEHURST, S. Potential of a Controllable Engine Cooling System to Reduce NOx Emissions in Diesel Engines. **SAE International**, p. 5, 2004.

PAUL, S.; KUMAR, S. Subsethood-product fuzzy neural inference system (SuPFuNIS). **IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council**, v. 13, n. 3, p. 578–99, 2002.

RABELO, A. Machine Learning: o que é e qual sua influência no marketing digital? Disponível em: <<https://marketingdeconteudo.com/machine-learning/>>. Acesso em: 23/9/2018.

RIO, R. L. Fan drive system [19]. , 1976. United States of America.

ROMANOV, V. A.; KHOZENIUK, N. A. Experience of the Diesel Engine Cooling System Simulation. **Procedia Engineering**, v. 150, p. 490–496, 2016. The Author(s).

SANDU, V. Experimental study on Diesel engine fitted with Visco fan drive. **Bulletim of the Transilvania University of Brasov**, v. 9, n. 1, p. 1–6, 2016.

SANTOS, E. M. **Teoria e Aplicação de Support Vector Machines à Aprendizagem e Reconhecimento de Objetos Baseado na Aparência**, 2002. Universidade Federal da Paraíba.

SCHÖLKOPF, B. The kernel trick for distances. **Advances in Neural Information Processing Systems 13**, v. 13, p. 301–307, 2001.

SCHOLKOPF, B.; SMOLA, A. J. **Learning with Kernels, Support Vector Machines, Regularization, Optimization and Beyond**. London, England: MIT Press, 2002.

SILVA RIBEIRO, D. J. **Support Vector Machines na Previsão do Comportamento de uma ETAR**, 2012. Universidade do Minho.

SMOLA, A. J.; SCHÖLKOPF, B. A tutorial on support vector regression. **Statistics and Computing**, v. 14, n. 3, p. 199–222, 2004.

SUYKENS, J. A. K.; VAN GESTEL, T.; DE BRABANTER, J.; DE MOOR, B.;

VANDEWALLE, J. **Least Squares Support Vector Machines**. Leuven, Belgium: World Scientific Publishing, 2002.

VAPNIK, V. N.; JORDAN, M.; F, J. **The Nature of Statistical Learning Theory**. 2nd ed. New York, USA: Springer Verrlag New York Inc., 1999.

ZENDEHBOUDI, S.; REZAEI, N.; LOHI, A. Applications of hybrid models in chemical, petroleum, and energy systems: A systematic review. **Applied Energy**, v. 228, n. December 2017, p. 2539–2566, 2018. Elsevier.

ZHIQIANG, C.; FEIYUE, M.; HUAYING, L.; XIANGBO, L.; SHOUSONG, H. The Hydraulic-Driven Fan Speed Control System Research Based on Fuzzy PID. **International Conference on Robots & Intelligent System (ICRIS)**, p. 444–448, 2016.

ZHOU, X., ZHANG, X., & WANG, B. *Online Support Vector Machine: A Survey. Harmony Search Algorithm: Proceedings of the 2nd International Conference on Harmony Search Algorithm (ICHSA2015)*, p. 269-278, 2016.