



Escola Politécnica – Universidade de São Paulo
Departamento de Engenharia Elétrica Sistemas
Eletrônicos

PSI2672 – Práticas em Reconhecimento de Padrões, Modelagem e Neurocomputação

APRESENTAÇÃO FINAL – 21/06/2016

PROF. EMILIO DEL MORAL HERNANDEZ

Integrantes do Grupo e Divisão de Tarefas



Ana Carolina Borg Ferreira -
8587517

- Seleção de Características
- Classificador
- Treinamento Redes Neurais



Lucas Rebelo Dal' Bello –
8582999

- Extração Características
- Classificador
- Treinamento Redes Neurais



Daniel Carlos de Melo
Marcilio – 8042801

- Coleta Dados Regressor
- Pré-Processamento Regressor



Pedro Dias Longhitano –
8988960

- Coleta Dados Regressor
- Pré-Processamento Regressor

Sumário

1. Reconhecimento de Movimentos com Acelerômetro Vestível

- Introdução e Motivação
- Coleta de Dados
- Pré-Processamento e Extração de Características
- Seleção de Características
- Treinamento da Rede
- Resultados
- Conclusão

2. Estimador do Índice de Qualidade da Água (IQA)

- Introdução e Motivação
- Coleta de Dados
- Pré-Processamento
- Treinamento da Rede
- Resultados
- Conclusão

Introdução e Motivação

Descrição	Reconhecimento de movimentos	Utilidade	Esportes	Medidas de desempenho
	Texas Instruments ez430-Chronos		Cuidados Médicos	Deteção de movimentos
	Acelerações nos eixos x, y e z		Segurança e Rastreamento	Monitoramento de atividades
			Residências Inteligentes	Interação com o Ambiente

Introdução e Motivação

ENTRADAS

Acelerações nos eixos x, y e z

Amostragem e Segmentação no Tempo

Extração de Características e Redução de Dimensionalidade

SAÍDAS

1 - Deitado

2 - Sentado

3 - Em Pé

4 - Andando

5 - Correndo

6 - Subindo Escada

7 - Descendo Escada

8 - Trabalhando no Computador

Coleta de Dados



Laboratório da Poli

- Sistema Supervisor para Aquisição de Dados
- Obtenção de Sinais de Diferentes Movimentações
- Desenvolvido por Mestrando da Poli Fernando Ginez da Silva em 2013

Pré-Processamento

Segmentação

- Amostragem a 33 Hz
- Tamanho da Janela: Compromisso entre Informação e Resolução

Filtragem

- Passa-baixas: remoção de ruídos e erros de amostragem
- Passa-altas: separar sinal da aceleração da gravidade

Extração de Características (SILVA, 2013)

DOMÍNIO DO TEMPO (19)

Correlação
Entre os Eixos
(xy, xz, yz)

Obliquidade
(x, y, z)

Curtose (x, y, z)

Área da
Magnitude do
Sinal

Média (x, y, z)

Desvio Padrão
(x, y, z)

Amplitude
Máxima (x, y, z)

DOMÍNIO DA FREQUÊNCIA (12)

Entropia
Espectral (x, y,
z)

Energia Média
do Sinal (x, y,
z)

Frequência de
Pico (x, y, z)

Potência na
Frequência de
Pico (x, y, z)

Seleção de Características – ReliefF

Conceito

- Estimar qualidade de features, a partir de sua capacidade de distinção inter-classes

K-Vizinhos Mais Próximos

Saídas

- Lista ranqueada de features
- Pesos de cada feature

Algorithm ReliefF

Input: for each training instance a vector of attribute values and the class value

Output: the vector W of estimations of the qualities of attributes

1. set all weights $W[A] := 0.0$;
2. **for** $i := 1$ **to** m **do begin**
3. randomly select an instance R_i ;
4. find k nearest hits H_j ;
5. **for** each class $C \neq class(R_i)$ **do**
6. from class C find k nearest misses $M_j(C)$;
7. **for** $A := 1$ **to** a **do**
8. $W[A] := W[A] - \sum_{j=1}^k \text{diff}(A, R_i, H_j) / (m \cdot k) +$
9. $\sum_{C \neq class(R_i)} \left[\frac{P(C)}{1 - P(class(R_i))} \sum_{j=1}^k \text{diff}(A, R_i, M_j(C)) \right] / (m \cdot k)$;
10. **end;**

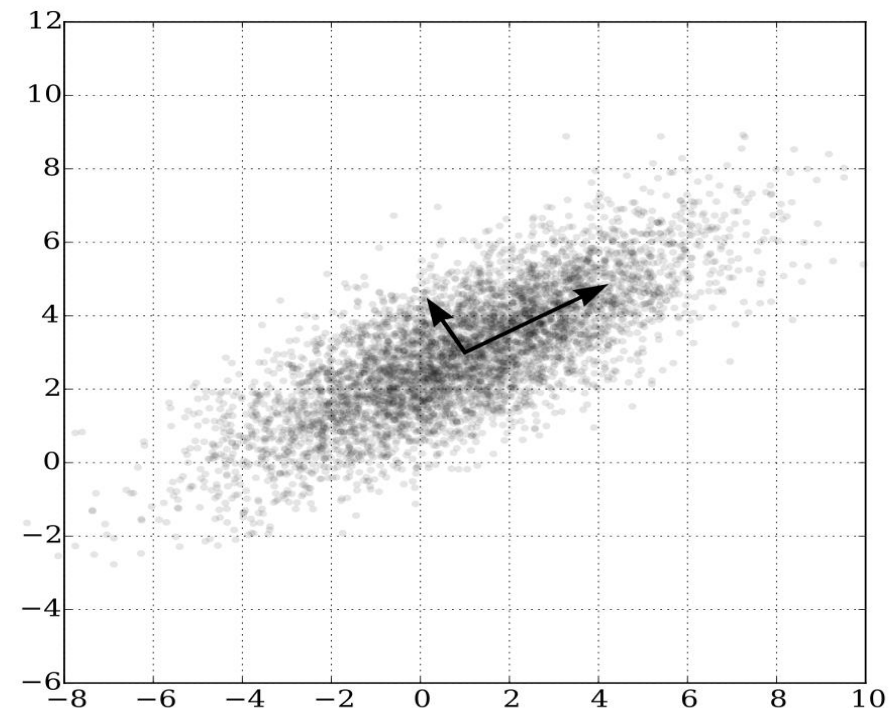
Fonte: (Sikonja & Kononenko, 2003)

Seleção de Características – PCA

Conceito

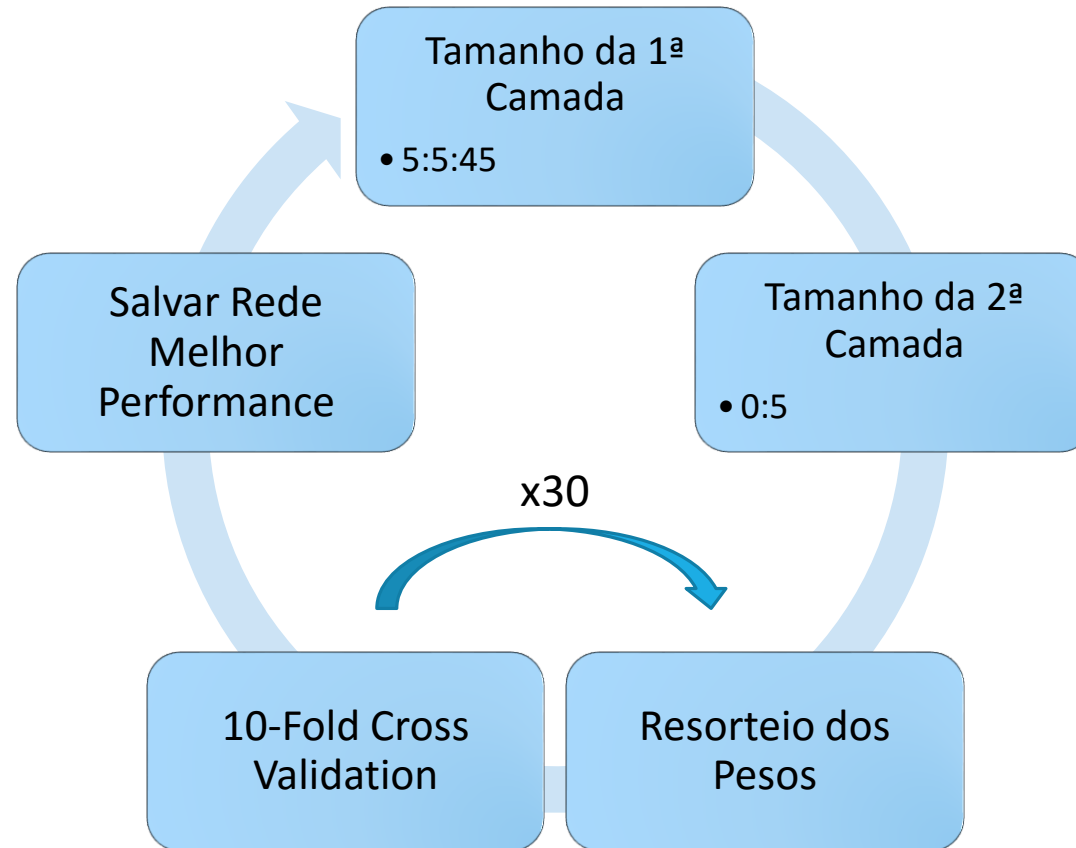
- Procedimento que usa transformação ortogonal para converter variáveis correlacionadas em componentes linearmente não-correlacionados

Seleção de Componentes com Maior Variância



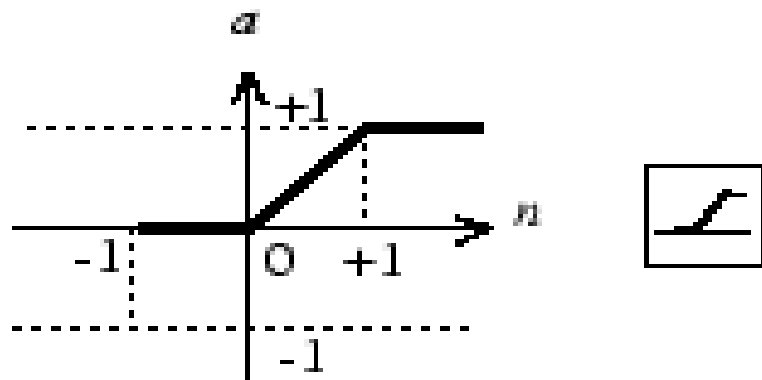
Fonte: Wikimedia Commons

Treinamento da Rede – Loops de Otimização



Treinamento da Rede – Camada de Saída

SATLIN

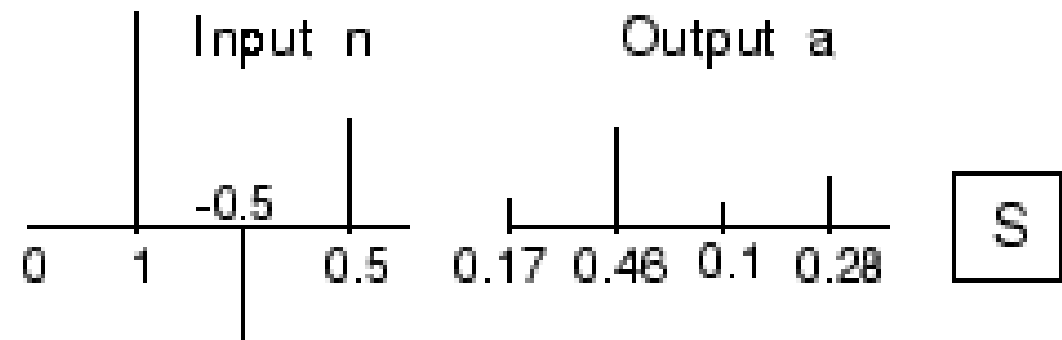


$$a = \text{satlin}(n)$$

Satlin Transfer Function

Fonte: Mathworks

SOFTMAX



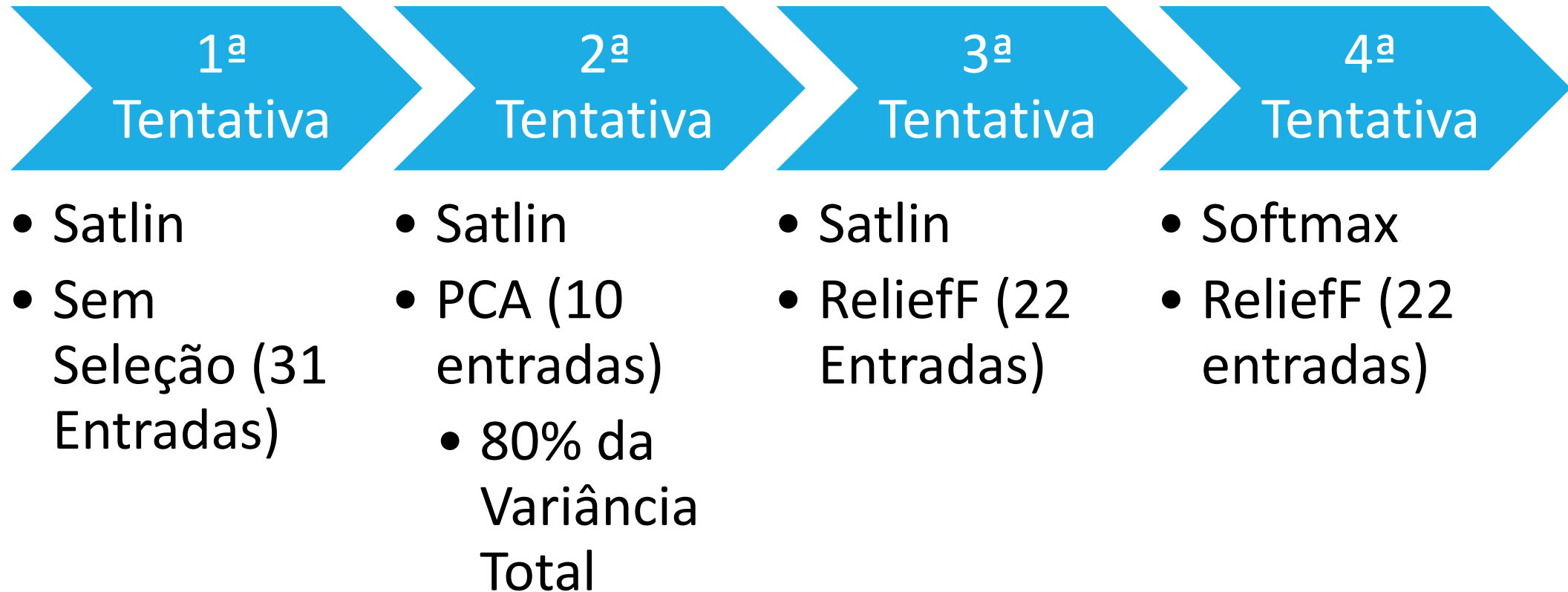
$$a = \text{softmax}(n)$$

Softmax Transfer Function

$$\text{Softmax}(n) = \frac{\exp(n)}{\sum(\exp(n))}$$

Fonte: Mathworks

Treinamento da Rede - Tentativas



Resultados

Confusion Matrix

	1	2	3	4	5	6	7	8	
1	198 8.0%	39 1.6%	47 1.9%	3 0.1%	1 0.0%	0 0.0%	2 0.1%	23 0.9%	63.3%
2	48 1.9%	308 12.5%	54 2.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	29 1.2%	70.2%
3	40 1.6%	41 1.7%	206 8.4%	1 0.0%	0 0.0%	2 0.1%	1 0.0%	23 0.9%	65.6%
4	7 0.3%	9 0.4%	10 0.4%	288 11.7%	0 0.0%	31 1.3%	9 0.4%	8 0.3%	79.6%
5	0 0.0%	0 0.0%	0 0.0%	4 0.2%	218 8.9%	0 0.0%	0 0.0%	0 0.0%	98.2%
6	1 0.0%	1 0.0%	1 0.0%	29 1.2%	1 0.0%	191 7.8%	17 0.7%	0 0.0%	79.3%
7	2 0.1%	0 0.0%	1 0.0%	2 0.1%	0 0.0%	18 0.7%	206 8.4%	0 0.0%	90.0%
8	20 0.8%	15 0.6%	15 0.6%	7 0.3%	0 0.0%	0 0.0%	0 0.0%	286 11.6%	83.4%
	62.7%	74.6%	61.7%	86.2%	99.1%	78.9%	87.7%	77.5%	77.2%
	37.3%	25.4%	38.3%	13.8%	0.9%	21.1%	12.3%	22.5%	22.8%
	1	2	3	4	5	6	7	8	
									Target Class

1ª Tentativa

Confusion Matrix

	1	2	3	4	5	6	7	8	
1	182 7.4%	52 2.1%	55 2.2%	4 0.2%	0 0.0%	3 0.1%	3 0.1%	21 0.9%	56.9%
2	50 2.0%	246 10.0%	52 2.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	20 0.8%	66.8%
3	47 1.9%	78 3.2%	184 7.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	38 1.5%	53.0%
4	10 0.4%	8 0.3%	6 0.2%	260 10.6%	0 0.0%	32 1.3%	5 0.2%	7 0.3%	79.3%
5	0 0.0%	0 0.0%	1 0.0%	5 0.2%	218 8.9%	1 0.0%	0 0.0%	0 0.0%	96.9%
6	1 0.0%	2 0.1%	6 0.2%	49 2.0%	2 0.1%	173 7.0%	23 0.9%	2 0.1%	67.1%
7	2 0.1%	1 0.0%	3 0.1%	6 0.2%	0 0.0%	33 1.3%	204 8.3%	0 0.0%	81.9%
8	24 1.0%	26 1.1%	27 1.1%	10 0.4%	0 0.0%	0 0.0%	0 0.0%	281 11.4%	76.4%
	57.6%	59.6%	55.1%	77.8%	99.1%	71.5%	86.8%	76.2%	71.0%
	42.4%	40.4%	44.9%	22.2%	0.9%	28.5%	13.2%	23.8%	29.0%
	1	2	3	4	5	6	7	8	
									Target Class

2ª Tentativa

Confusion Matrix

	1	2	3	4	5	6	7	8	
1	179 7.3%	43 1.7%	38 1.5%	3 0.1%	0 0.0%	0 0.0%	2 0.1%	12 0.5%	64.6%
2	53 2.2%	291 11.8%	47 1.9%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	15 0.6%	71.7%
3	46 1.9%	55 2.2%	219 8.9%	2 0.1%	0 0.0%	1 0.0%	2 0.1%	25 1.0%	62.6%
4	8 0.3%	9 0.4%	4 0.2%	286 11.6%	0 0.0%	24 1.0%	4 0.2%	4 0.2%	84.4%
5	0 0.0%	0 0.0%	1 0.0%	4 0.2%	219 8.9%	0 0.0%	0 0.0%	0 0.0%	97.8%
6	1 0.0%	1 0.0%	3 0.1%	30 1.2%	1 0.0%	198 8.0%	23 0.9%	0 0.0%	77.0%
7	1 0.0%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	19 0.8%	203 8.2%	0 0.0%	90.6%
8	28 1.1%	14 0.6%	22 0.9%	8 0.3%	0 0.0%	0 0.0%	1 0.0%	313 12.7%	81.1%
	56.6%	70.5%	65.6%	85.6%	99.5%	81.8%	86.4%	84.8%	77.5%
	43.4%	29.5%	34.4%	14.4%	0.5%	18.2%	13.6%	15.2%	22.5%
	1	2	3	4	5	6	7	8	
									Target Class

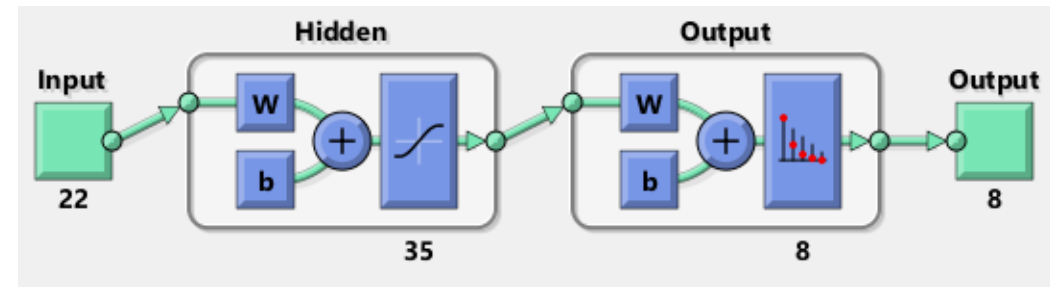
3ª Tentativa

Resultado Final

Confusion Matrix

Output Class \ Target Class	1	2	3	4	5	6	7	8	Accuracy
1	195 7.9%	37 1.5%	38 1.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	12 0.5%	69.1% 30.9%
2	49 2.0%	314 12.7%	56 2.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	16 0.6%	72.2% 27.8%
3	39 1.6%	37 1.5%	214 8.7%	1 0.0%	0 0.0%	0 0.0%	1 0.0%	11 0.4%	70.6% 29.4%
4	7 0.3%	9 0.4%	9 0.4%	302 12.3%	1 0.0%	25 1.0%	9 0.4%	3 0.1%	82.7% 17.3%
5	0 0.0%	0 0.0%	1 0.0%	4 0.2%	218 8.9%	0 0.0%	0 0.0%	0 0.0%	97.8% 2.2%
6	1 0.0%	2 0.1%	2 0.1%	18 0.7%	0 0.0%	200 8.1%	14 0.6%	0 0.0%	84.4% 15.6%
7	1 0.0%	0 0.0%	0 0.0%	3 0.1%	1 0.0%	17 0.7%	211 8.6%	0 0.0%	90.6% 9.4%
8	24 1.0%	14 0.6%	14 0.6%	6 0.2%	0 0.0%	0 0.0%	0 0.0%	327 13.3%	84.9% 15.1%
Overall	61.7% 38.3%	76.0% 24.0%	64.1% 35.9%	90.4% 9.6%	99.1% 0.9%	82.6% 17.4%	89.8% 10.2%	88.6% 11.4%	80.4% 19.6%

4ª Tentativa



Softmax

ReliefF

22 Entradas

Taxa de Acerto Total: 80,4%

Resultado Final - Entradas Utilizadas

1. Desvio Padrão y	2. Amplitude Máxima y	3. Desvio Padrão x	4. Área da Magnitude do Sinal	5. Amplitude Máxima x	6. Potência na Frequência de Pico y
7. Energia Média x	8. Potência na Frequência de Pico x	9. Energia Média y	10. Amplitude Máxima z	11. Desvio Padrão z	12. Frequência Pico y
13. Potência na Frequência de Pico z	14. Entropia Espectral y	15. Correlação xy	16. Frequência de Pico z	17. Frequência de Pico x	18. Entropia Espectral x
	19. Correlação xz	20. Correlação yz	21. Entropia Espectral z	22. Energia Média z	

Conclusões

Classificação Pouco Eficaz para Classes Estáticas

- Erro de cerca de 30% para essas classes
- Pouca distinção inter-classe

Boa Classificação de Movimentos

ReliefF Parece ter Eliminado Características Desnecessárias

Maiores Dificuldades

- Tempo de Processamento
- Com mais tempo, aplicar PCA sobre a configuração atual

Introdução e Motivação

Descrição Estimação da qualidade da água a partir de suas características bioquímicas

Baseado no IQA

Utilidade Adaptabilidade

Redução de Tempo e Custos

Introdução e Motivação

ENTRADAS

Oxigênio Dissolvido	pH	Demanda Bioquímica de Oxigênio (DBO)
Coliformes Fecais	Temperatura da Água	Nitrogênio Total
Fósforo Total	Turbidez	Sólidos Totais

SAÍDAS

IQA

- Variável Contínua
- Valores entre 0 e 100
- Faixas de Valores Definem a Qualidade

Coleta de Dados

Relatório da CETESB (Companhia Ambiental do Estado de São Paulo)

Qualidade das Águas Superficiais no Estado de São Paulo

Dezembro de 2014

Resultado dos Parâmetros e Indicadores de qualidade das águas

Código do Ponto: 65 - 0 0 SP 05 714 CP1V 02 100

Classe : 2

impresso em: 16/03/2015

Sistema Hídrico: Rio Capivari

Localização: No condomínio São Joaquim, em Vinhedo. - VINHEDO - SP

Tipo de Parametro: Campo								
Descrição do Parâmetro	Und.	Padrão CONAMA	05/02/2014 12:06:00	07/04/2014 14:00:00	09/06/2014 14:44:00	11/08/2014 18:10:00	08/10/2014 12:20:00	01/12/2014 15:11:00
Chuvas nas últimas 24h	-		Não	Não	Não	Não	Não	Não
Coloração	-		Cinza	Cinza	Cinza	Cinza	Cinza	Cinza
Condutividade	µS/cm		361	450	396	628	598	370
Oxigênio Dissolvido	mg/L	> 5	* < 0,4	* < 0,4	* < 0,4	* < 0,4	* < 0,4	* < 0,4
pH	U.pH	6 até 9	7,1	7	6,9	7,2	7	6,9
Temperatura da Água	°C		26	25	19,8	19,3	22,5	25,2
Temperatura do Ar	°C		31	31,5	26,5	24	28,1	34,8
Tipo de Parametro: Físicos e Químicos								
Descrição do Parâmetro	Und.	Padrão CONAMA	05/02/2014 12:06:00	07/04/2014 14:00:00	09/06/2014 14:44:00	11/08/2014 18:10:00	08/10/2014 12:20:00	01/12/2014 15:11:00
Alumínio Dissolvido	mg/L	< 0,1	< 0,1		< 0,1	< 0,1		< 0,1
Alumínio Total	mg/L		0,4		0,7	0,4		1
Bário Total	mg/L	< 0,7	0,09		0,1	0,1		0,1
Cádmio Total	mg/L	< 0,001	< 0,001		< 0,001	< 0,001		< 0,001
Carbono Orgânico Total	mg/L		23,9	37,5	38,4	55	64,8	23,9
Chumbo Total	mg/L	< 0,01	< 0,01		< 0,01	< 0,01		< 0,01
Cloreto Total	mg/L	< 250	28	57	29	82	48	33
Cobre Dissolvido	mg/L	< 0,009	< 0,005		< 0,005	< 0,005		< 0,005
Cobre Total	mg/L		< 0,005		0,01	0,007		< 0,005
Cromo Total	mg/L	< 0,05	< 0,05		< 0,05	< 0,05		< 0,05

Pré-Processamento

Valores Ausentes

Substituição pela Média

Outliers

Comparação com Múltiplos da Mediana

Substituição pelo Máximo ou Mínimo não-Outlier

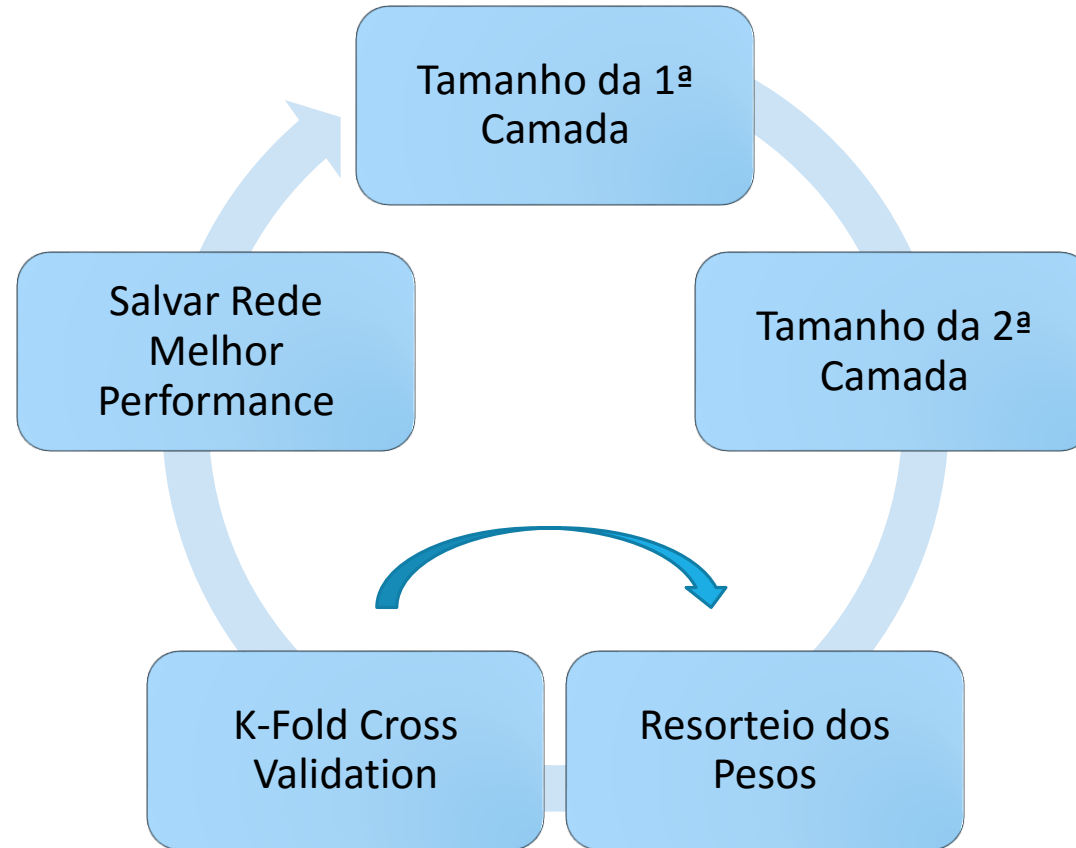
Valores Não
Atômicos

Substituição pelo Máximo ou Mínimo do Intervalo

Desbalanceamento

Divisão das Faixas de Valores

Treinamento da Rede – Loops de Otimização



Treinamento da Rede – Algoritmos de Aprendizado

LEVENBERG-MARQUARDT BACKPROPAGATION

SCALED CONJUGATE GRADIENT
BACKPROPAGATION

Função de Performance: Erro
Quadrático Médio

Função de Performance: Erro
Absoluto Médio

Utiliza Mais Memória

Utiliza Menos Memória

Leva Menos Tempo

Leva Mais Tempo

Fonte: Mathworks

Treinamento da Rede - Tentativas



1ª Tentativa

- Levenberg-Marquardt
- Tamanho da 1ª Camada: 2:2:16
- Tamanho da 2ª Camada: 0:5
- 5-Fold Cross Validation
- 30 Ressorteios

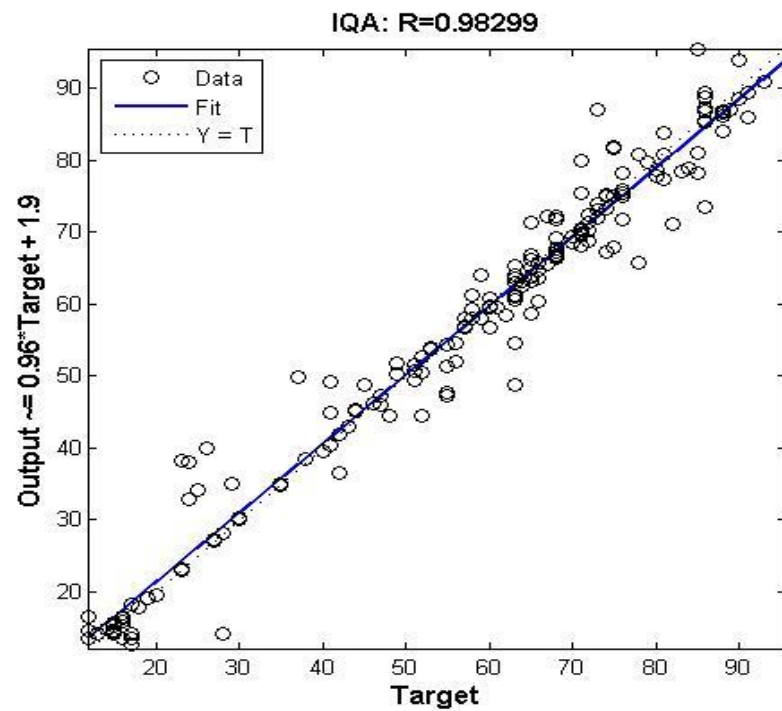


2ª Tentativa

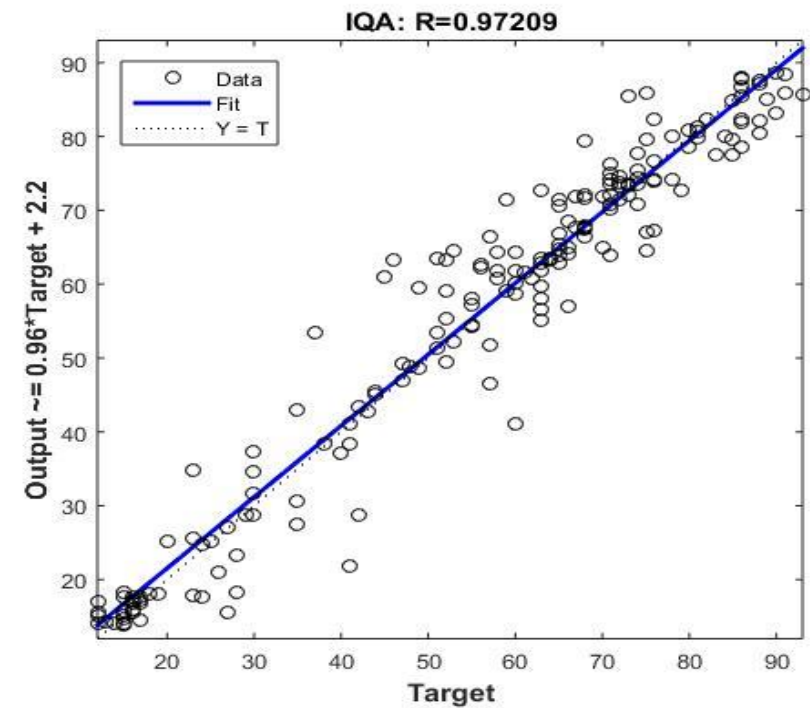
- Scaled Conjugate Gradient
- Tamanho da 1ª Camada: 4:2:16
- Tamanho da 2ª Camada: 0:3
- 10-Fold Cross Validation
- 40 Ressorteios

Resultados

TENTATIVA 1



TENTATIVA 2



Resultados

TENTATIVA 1

TENTATIVA 2

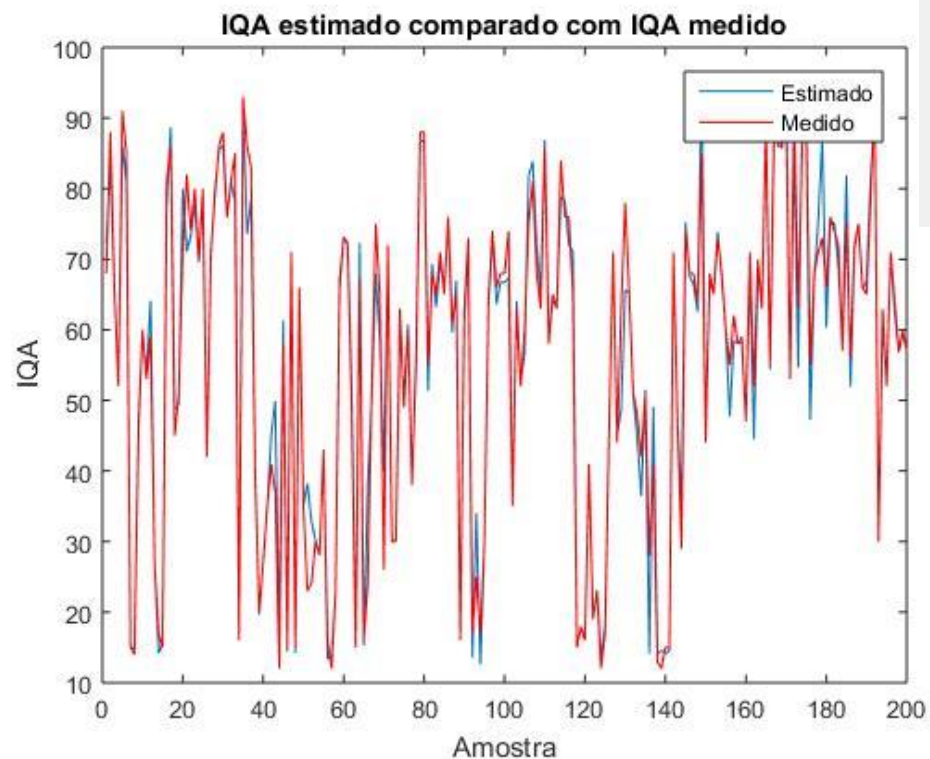
**Máximo Módulo
do Erro: 66,3%**

**Máximo Módulo
do Erro: 51,1%**

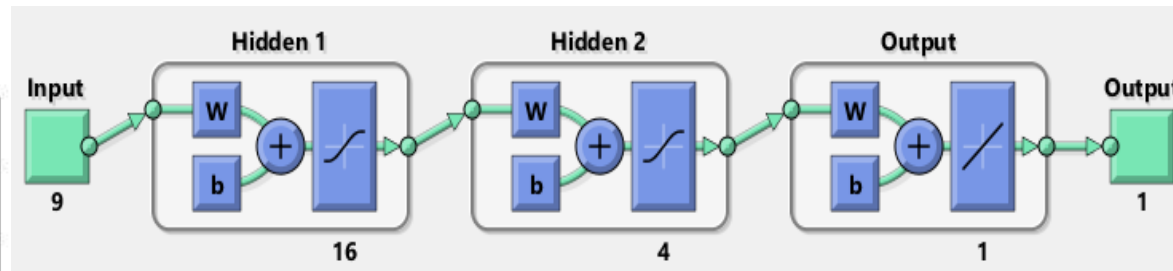
**Média do Módulo
do Erro: 5,9%**

**Média do Módulo
do Erro: 8,1%**

Resultado Final



1ª Tentativa



Levenberg-Marquardt

Tamanho da 1ª Camada: 2:2:16

Tamanho da 2ª Camada: 0:5

5-Fold Cross Validation

30 Ressorteios

Conclusões

Erros máximos elevados indicam imprecisões nos dados de entrada

Pré-Processamento pode ser revisitado

Para o futuro, treinar com seleção de características (PCA)

Categoria	Ponderação
ÓTIMA	$79 < IQA \leq 100$
BOA	$51 < IQA \leq 79$
REGULAR	$36 < IQA \leq 51$
RUIM	$19 < IQA \leq 36$
PÉSSIMA	$IQA \leq 19$

Um erro médio de 6% não é suficiente para deslocar o ponto médio entre as faixas de qualidade

Bibliografia

Silva, F. G. (2013). *Reconhecimento de movimentos humanos utilizando um acelerômetro e inteligência computacional*. São Paulo.

Sikonja & Kononenko (2003). Theoretical and Empirical Analysis of ReliefF and RReliefF

CETESB. (2014). *Qualidade das Águas Superficiais no Estado de São Paulo*.

CETESB. (s.d.). *Apêndice C - Índices de Qualidade das Águas*.

Filho, A. C. (2006). *Análise de qualidade de água utilizando rede neural*.
