

# PSI2672

## PRÁTICAS EM RECONHECIMENTO DE PADRÕES, MODELAGEM E NEUROCOMPUTAÇÃO

ALUNOS:	André V. O. Maggio	Nº:8494372
	Marcos Henrique Carvalho Silva	Nº:7630920
	Matheus Barros Manini	Nº:7211036
PROFESSOR:	Emilio Del Moral Hernandez	

Escola Politécnica da Universidade de São Paulo Desde 1893



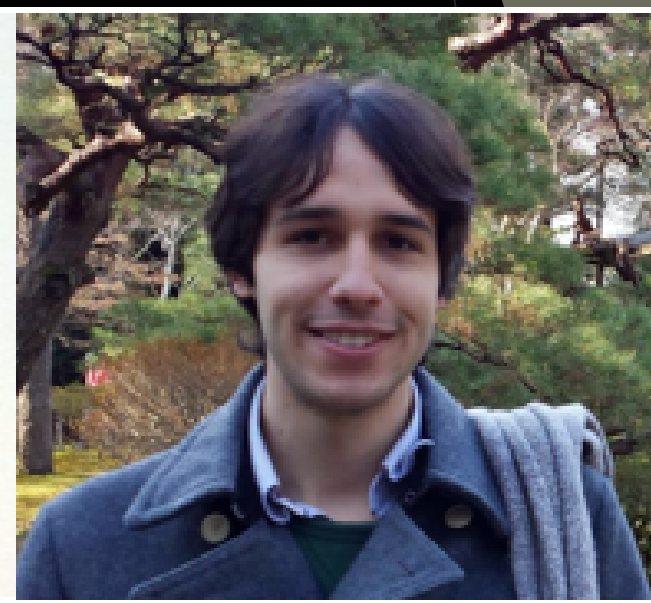
A melhor escola de engenharia da América Latina  
segundo o Times Higher Education em 2014



André Vinícius  
Oliveira Maggio



Marcos Henrique  
Carvalho Silva



Matheus Barros  
Manini

Identificação de Fraudes em  
Compras de Cartão

Reconhecimento de Padrões nos  
Sobreviventes e Vítimas do Titanic

Estimativa da dose de pigmentos  
para tintas a partir de uma amostra

# ORGANIZAÇÃO

- ▶ Cada membro tem conhecimento sobre todos projetos e tem como função opinar, corrigir e fazer sugestões sobre os projetos dos outros "especialistas".
- ▶ É frequente que um integrante do grupo influencie o projeto do outro com ideias novas, assim como contribua com dados. É frequente também que um integrante do grupo faça tarefas relacionada a um projeto pelo qual ele não é principalmente responsável.
- ▶ O aluno Matheus está principalmente responsável pelo projeto de "Fraudes em Cartões", o aluno André pelo projeto de "Pigmentos" e o aluno Marcos pelo projeto de "Vítimas do Titanic".

# Exemplos de integração dos membros nos diversos projetos são:

- ▶ Houve troca do projeto do Marcos de "Cobras Venenosas" para "Vítimas do Titanic" por sugestão do aluno Matheus. Este sugeriu o tema, além de apresentar ao Marcos toda a base de dados dos tripulantes do Titanic. Auxiliou o Marcos também na escolha do vetor X.
- ▶ Marcos cedeu suas notas fiscais ao Matheus para auxilia-lo no projeto de detecção de fraude.
- ▶ André é geralmente o responsável por montar o relatório final, organizando as partes de cada membro.
- ▶ Existe ampla comunicação entre os integrantes do grupo, de forma que cada um tem inteiro conhecimento daquilo que está sendo executado por outro membro, além de que cada um revisa o relatório do outro.
- ▶ O André gerou o código C do projeto Vítimas do Titanic para a demonstração. O André também idealizou a demonstração do projeto Vítimas do Titanic.

# Identificação de Fraudes em Compras de Cartão



# Suposições

- ▶ Quaisquer meios de compras (online ou fisicamente)
- ▶ Comportamento como fator determinante
- ▶ Tamanho distinto das classes de classificação foi desconsiderado
- ▶ Desconsiderado especificação do algoritmo em tipos deferentes de fraudes (como roubo, perda de cartão, clonagem e etc)

# O que define comportamento?

- ▶ Rotinas: horário de almoço, horário de aulas/trabalho
  - ▶ Dia da semana
  - ▶ Horário
- ▶ Localização das lojas
  - ▶ Estado de onde compra produtos (serve como identificação de compras online)
- ▶ Tipo de produto/loja
  - ▶ Preferências por almoçar fora, ou comprar muitos remédios, etc
- ▶ Quanto se gasta, em quantas vezes se paga

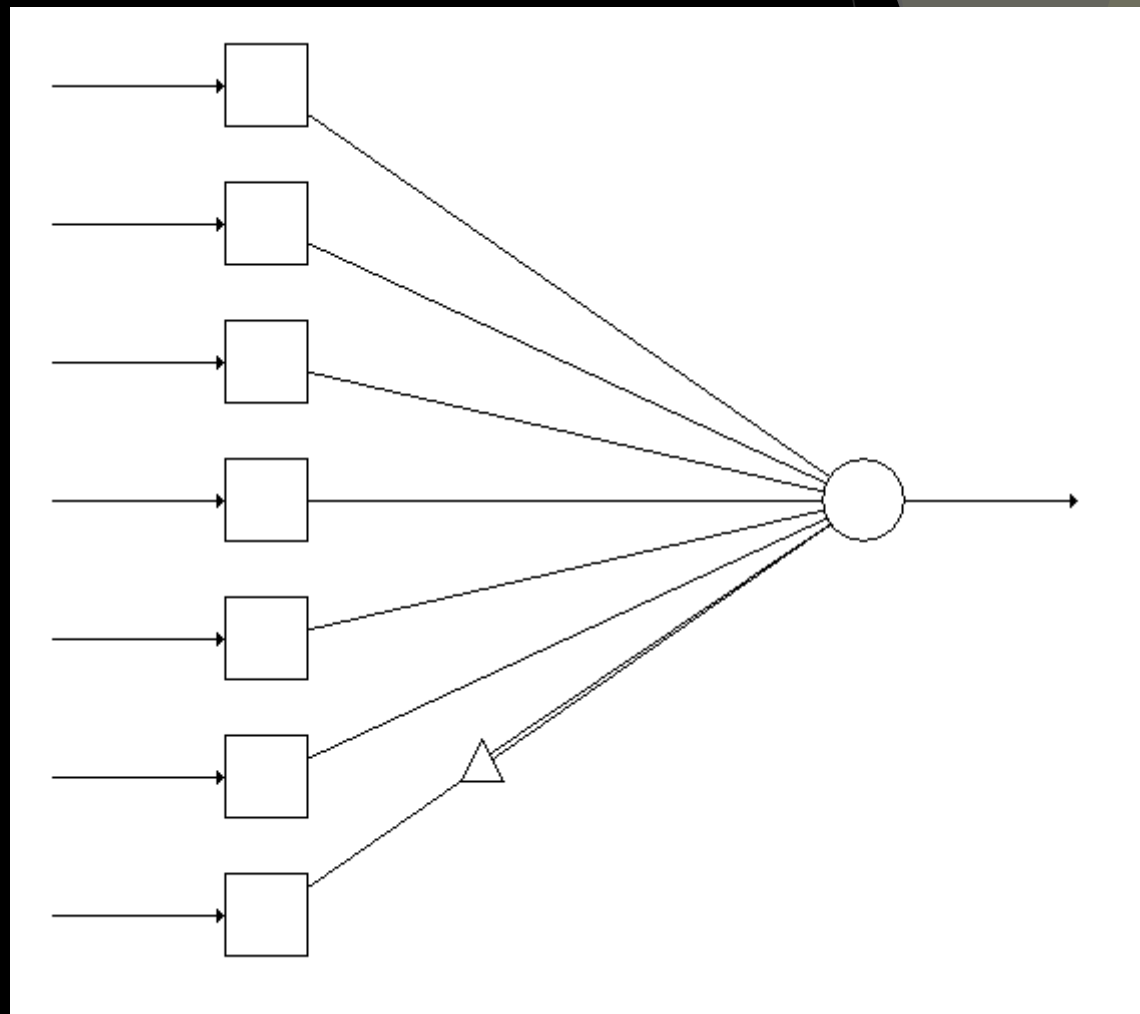
# Variáveis consideradas

- ▶ Estado do Utilizador do Cartão
- ▶ Estado da Loja
- ▶ Tipo de Loja
- ▶ Custo da Compra
- ▶ Dia da Semana
- ▶ Horário
- ▶ Forma de Pagamento

Tudo isto está identificado na nota fiscal de compra!

# Rede Neural Inicial

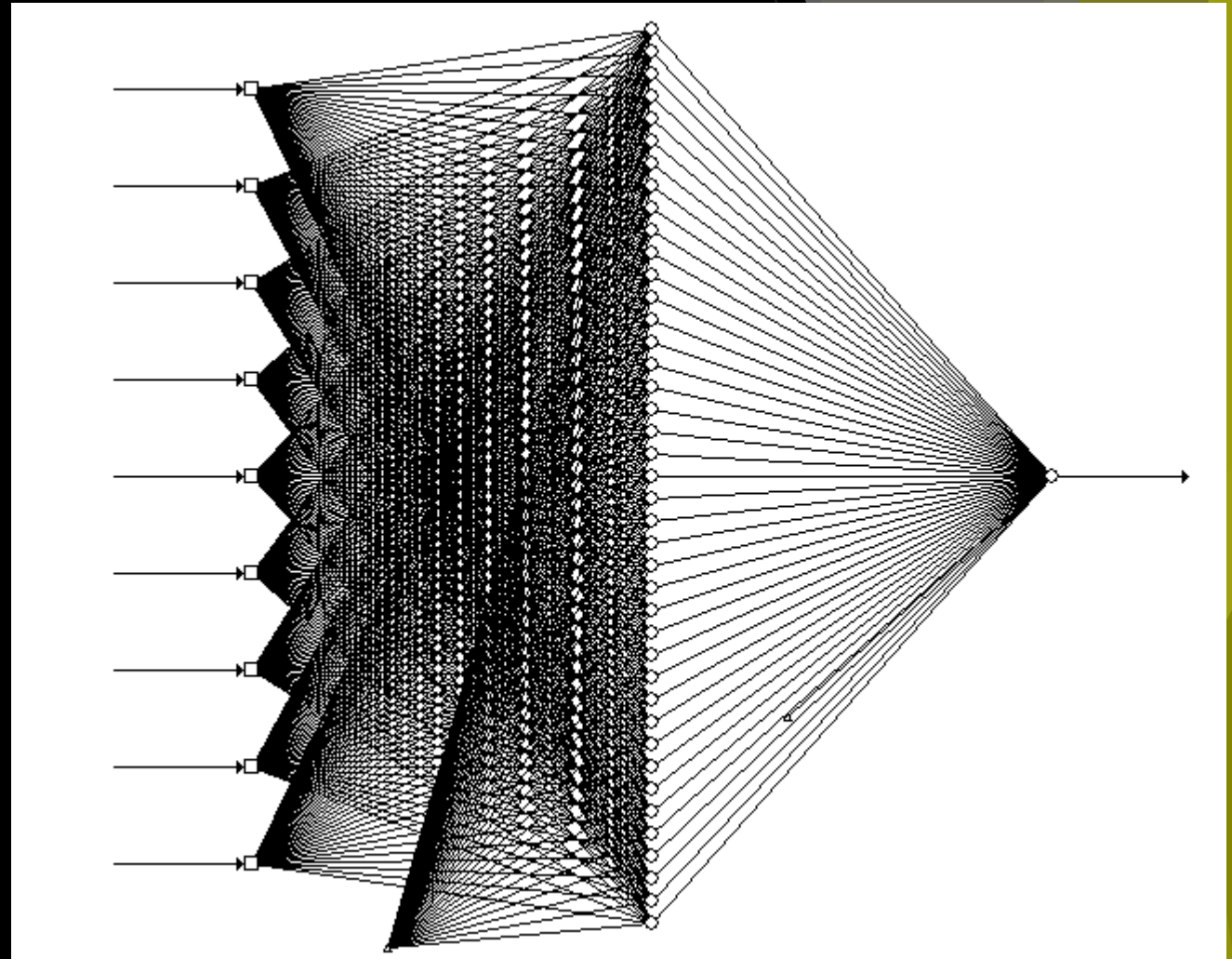
- ▶ Rede 7-1
- ▶ Neurônio único: regressor linear
- ▶ Técnica amplamente utilizada em negócios
- ▶ Classificação muito simplória, apenas um hiperplano linear



# Proposta

- ▶ Rede 9-41-1
- ▶ Último neurônio tem característica digital, ideal para identificação de padrões
- ▶ Mas por que 9 entradas? Por que 41 neurônios?

Chute!



# Chute de Engenharia

- ▶ Utilizando o MATLAB pode-se testar redes com duas camadas de modo exaustivo, treinando redes de 1 até 50 neurônios
- ▶ O final da série de testes apresenta aparência monotonica, portanto foi considerado overlearning para mais de 50 neurônios
- ▶ Chegou-se ao resultado 41

```
melhor = 0;
melhorp = 1;

x = input';
t = output';

for i=1:50

    % Create a Pattern Recognition Network
    net = patternnet(i);

    % Setup Division of Data for Training, Validation, Testing
    net.divideParam.trainRatio = 50/100;
    net.divideParam.valRatio = 0/100;
    net.divideParam.testRatio = 50/100;
    net.trainParam.epochs = 100;

    % Train the Networka
    [net,tr] = train(net,x,t);

    % Test the Network
    y = net(x);
    performance = perform(net,t,y)

    if(performance < melhorp)
        melhor = i;
        melhorp = performance;
    end

end
```

# Chute de Engenharia

- ▶ Utilizou-se variáveis de controle mais variáveis descritas resultando em 9 entradas.
- ▶ Para um número qualquer de neurônios (utilizou-se 5 e 41) removeu-se uma das variáveis e treinou-se a rede.
- ▶ Nenhuma redução de variável apresentou aumento na performance, e portanto manteve-se as 9.

```
% Test the Network
y = net(x);
performance = perform(net,t,y)

melhor = 0;
melhorp = performance;

for i=1:size(input,2)

    x = input';
    x(i,:) = [];

    net = patternnet(5);

    % Setup Division of Data for Training, Validation, Testing
    net.divideParam.trainRatio = 50/100;
    net.divideParam.valRatio = 0/100;
    net.divideParam.testRatio = 25/100;
    net.trainParam.epochs = 100;

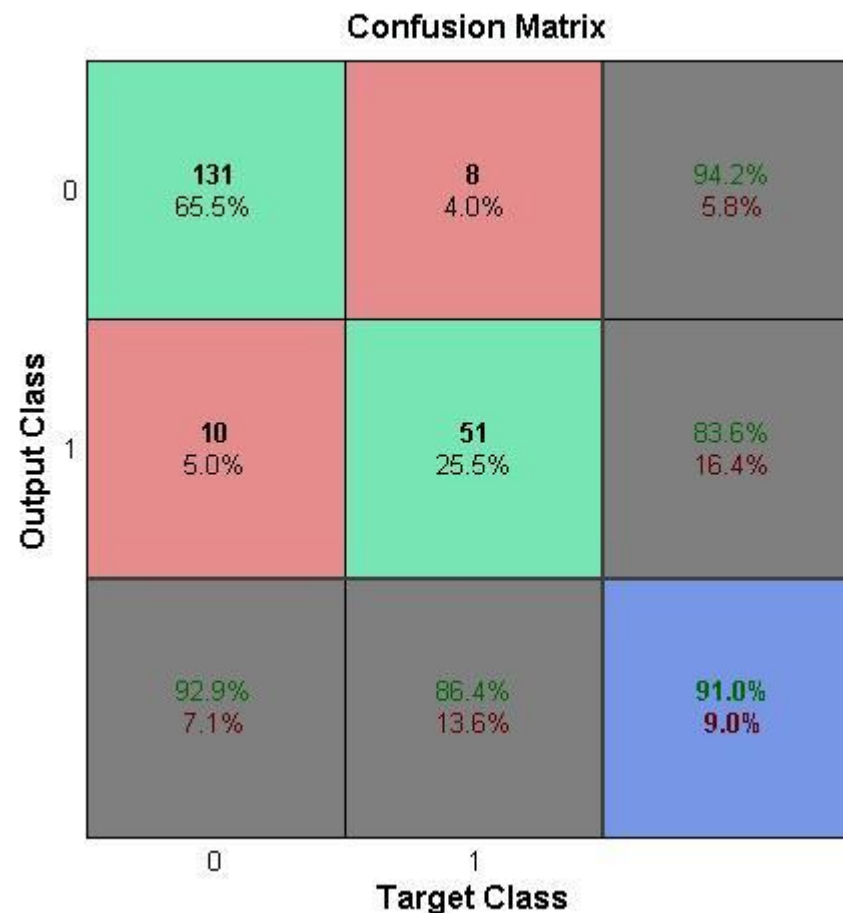
    % Train the Networka
    [net,tr] = train(net,x,t);

    % Test the Network
    y = net(x);
    performance = perform(net,t,y)

    if(performance < melhorp)
        melhor = i;
        melhorp = performance;
    end
```

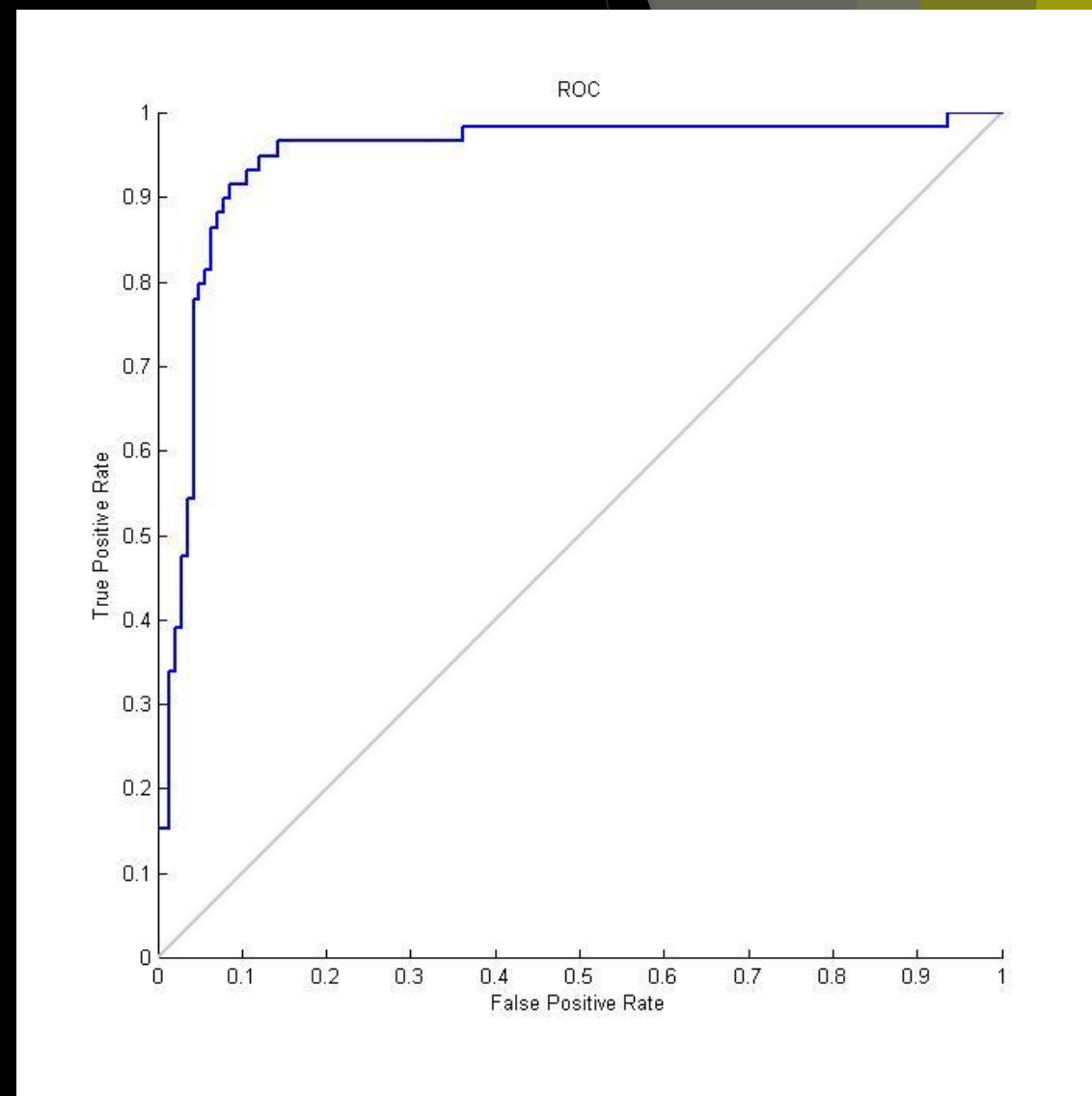
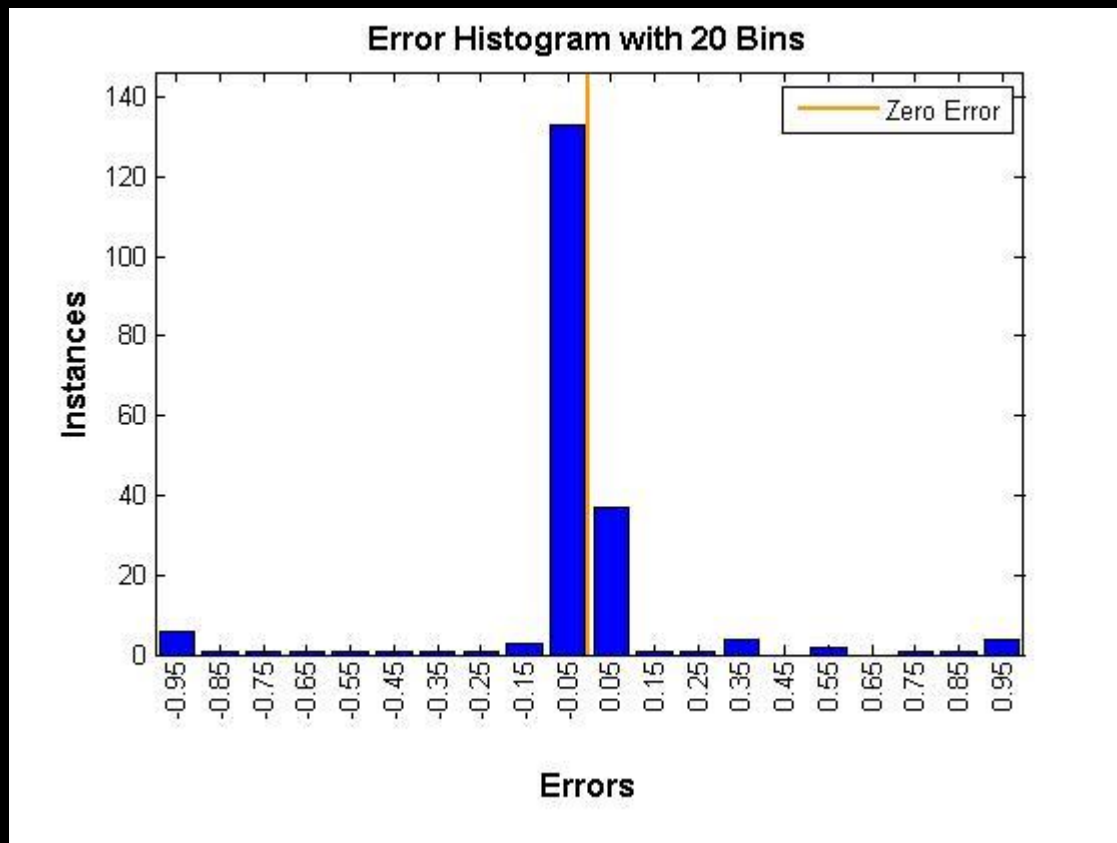
# Resultados - Matriz de Confusão

- ▶ Alta taxa de acerto
- ▶ Erro varia entre 9~11%, dependendo do grupo de dados escolhido.
- ▶ Uso em conjunto com outros métodos: taxas ainda melhores!





# Resultados - ROC e Histograma



# Reconhecimento de Padrões nos Sobreviventes e Vítimas do Titanic

Projeto procurará definir fatores que influenciam na chance de sobrevivência num naufrágio

- ▶ Fatores Biológicos
- ▶ Fatores Sociais
- ▶ Fatores Geográficos
- ▶ Informação Privilegiada

# Artigo: Environmental Risks and Environmental Justice, Or How Titanic Risks Are Not So Titanic After All

Margarita V. Alario e William R. Freudenburg

## Artigo: Environmental Risks and Environmental Justice, Or How Titanic Risks Are Not So Titanic After All

- ▶ "the fact that—as was the case for the majority of the victims on the Titanic—actual risks are related to victims' socioeconomic as well as sociogeographic locations"

## Artigo: Environmental Risks and Environmental Justice, Or How Titanic Risks Are Not So Titanic After All

- ▶ "The first-class passengers, who included some of the richest and most prominent people in the world, were notified of the accident within the hour, and they began to be evacuated soon after. Those in first class who did not survive were largely those who decided to stay, in many cases because the Titanic was thought to be unsinkable"

# Artigo: Environmental Risks and Environmental Justice, Or How Titanic Risks Are Not So Titanic After All

- ▶ "Of more than 2200 passengers on board, 712 were saved, but only about 50 of them were traveling in third-class sections of the ship."

## Artigo: Environmental Risks and Environmental Justice, Or How Titanic Risks Are Not So Titanic After All

- ▶ "Survivors also included 70 in second-class compartments (...) By contrast, even though the ship carried fewer first-class passengers, nearly 600 of them were rescued."



# Artigo: Surviving the Titanic disaster: economic, natural and social determinants

Bruno S. Frey, David A. Savage e Benno Torgler

## Artigo: Surviving the Titanic disaster: economic, natural and social determinants

- ▶ "Our paper analyzes the determinants of who is more likely to survive such a tragic event."

## Artigo: Surviving the Titanic disaster: economic, natural and social determinants

- ▶ "(...) the Titanic disaster is so relevant in a more general context. It allows us to analyze behavior under extraordinary conditions, namely in a life and death situation."

## Artigo: Surviving the Titanic disaster: economic, natural and social determinants

- ▶ "The crew had access to more informational and relational resources and therefore had a higher survival chance than the passengers, in particular, the deck crew."

## Artigo: Surviving the Titanic disaster: economic, natural and social determinants

- ▶ "We find that people in their prime (ages 15-35) had a higher chance of survival than older people. Women of reproductive age and women with children also had a higher probability of being rescued".

# Artigo: Noblesse oblige? Determinants of survival in a life-and-death situation

Bruno S. Frey, David A. Savage e Benno Torgler

## Artigo: Noblesse oblige? Determinants of survival in a life-and-death situation

- ▶ "(...)the sinking of the Titanic. This dataset not only allows us to explore the behavioral consequences of an extraordinary event, but also provides evidence of an individual's reaction in a situation where there exists an excess of demand due to the shortage of lifeboats."

## Artigo: Noblesse oblige? Determinants of survival in a life-and-death situation

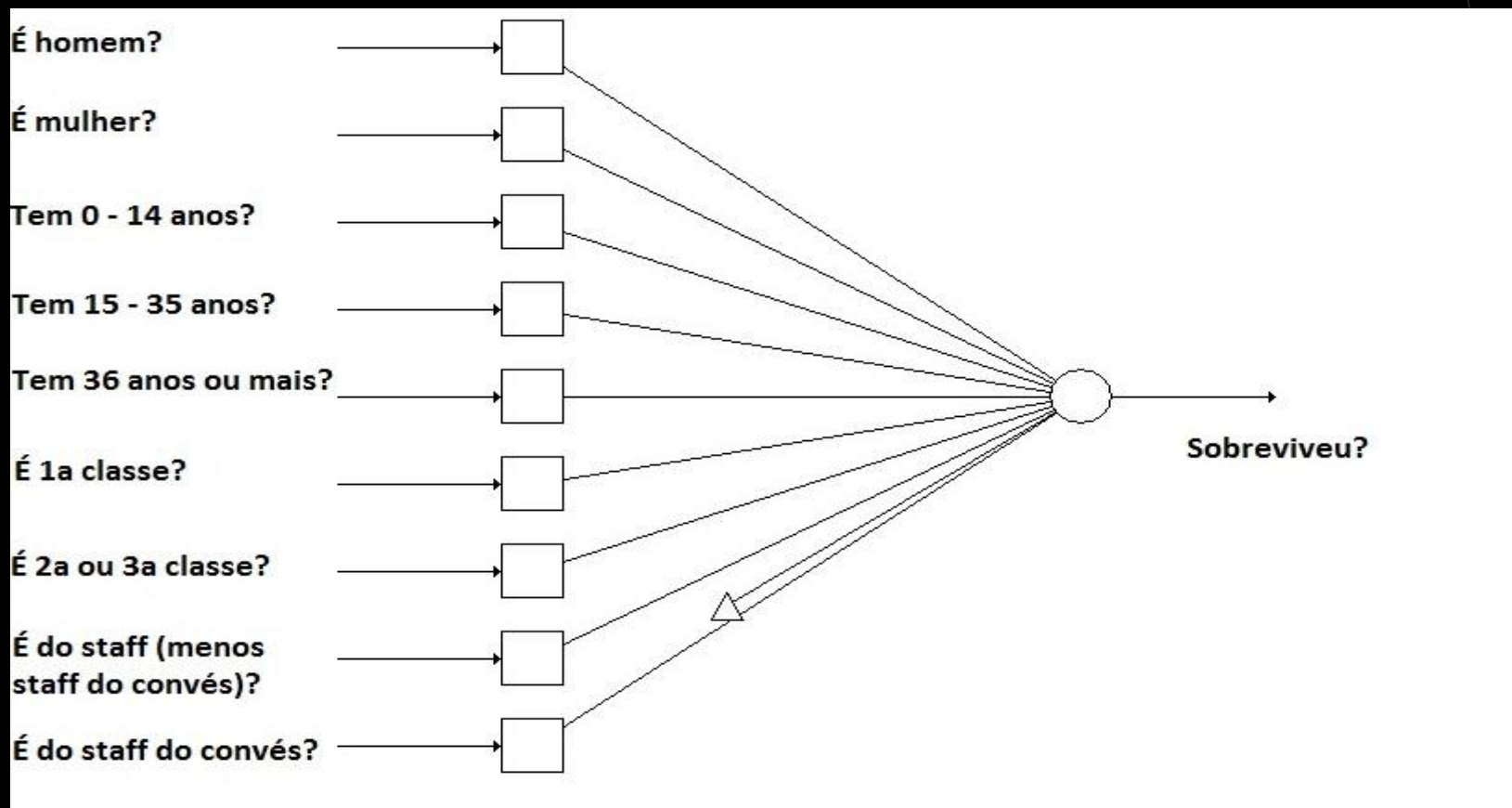
- "(...)altruism matter. The adherence to the norm "women and children first" is apparent in such a life-and-death situation. Being female rather than male increases the probability of survival between 23.7 percent and 53.9 percent, depending on the specification used. This is a large quantitative effect."

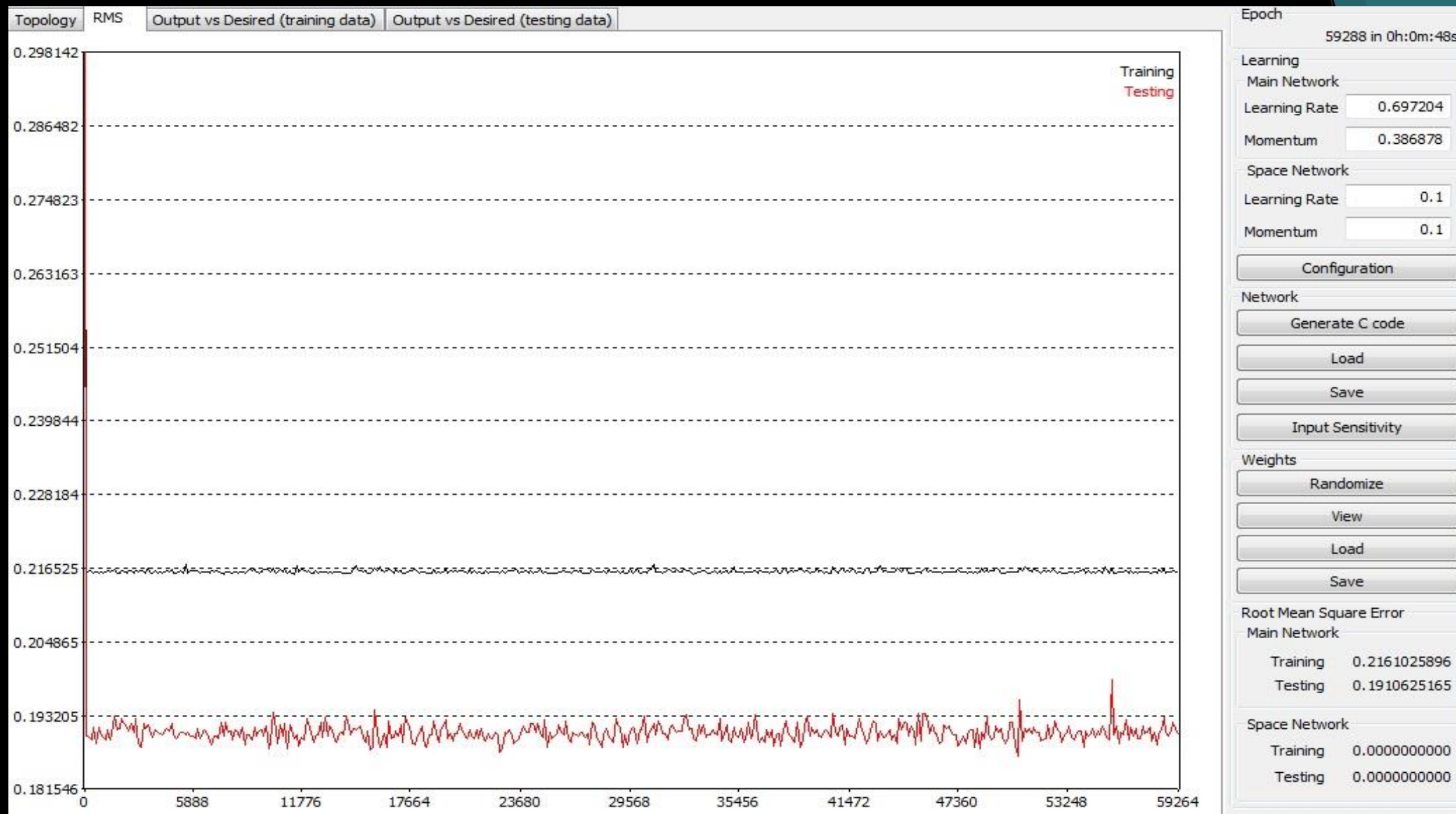


## Artigo: Noblesse oblige? Determinants of survival in a life-and-death situation

- ▶ "Similarly, being a child rather than a person of AGE 51+ (reference group) increases the probability of survival by about 30 percent. Having a child and being of reproductive age strongly raises the survival probability. Having a child also increases the probability of survival when males are considered also. Such results are in line with socio-biological theories"

# Rede Neural 9 -1



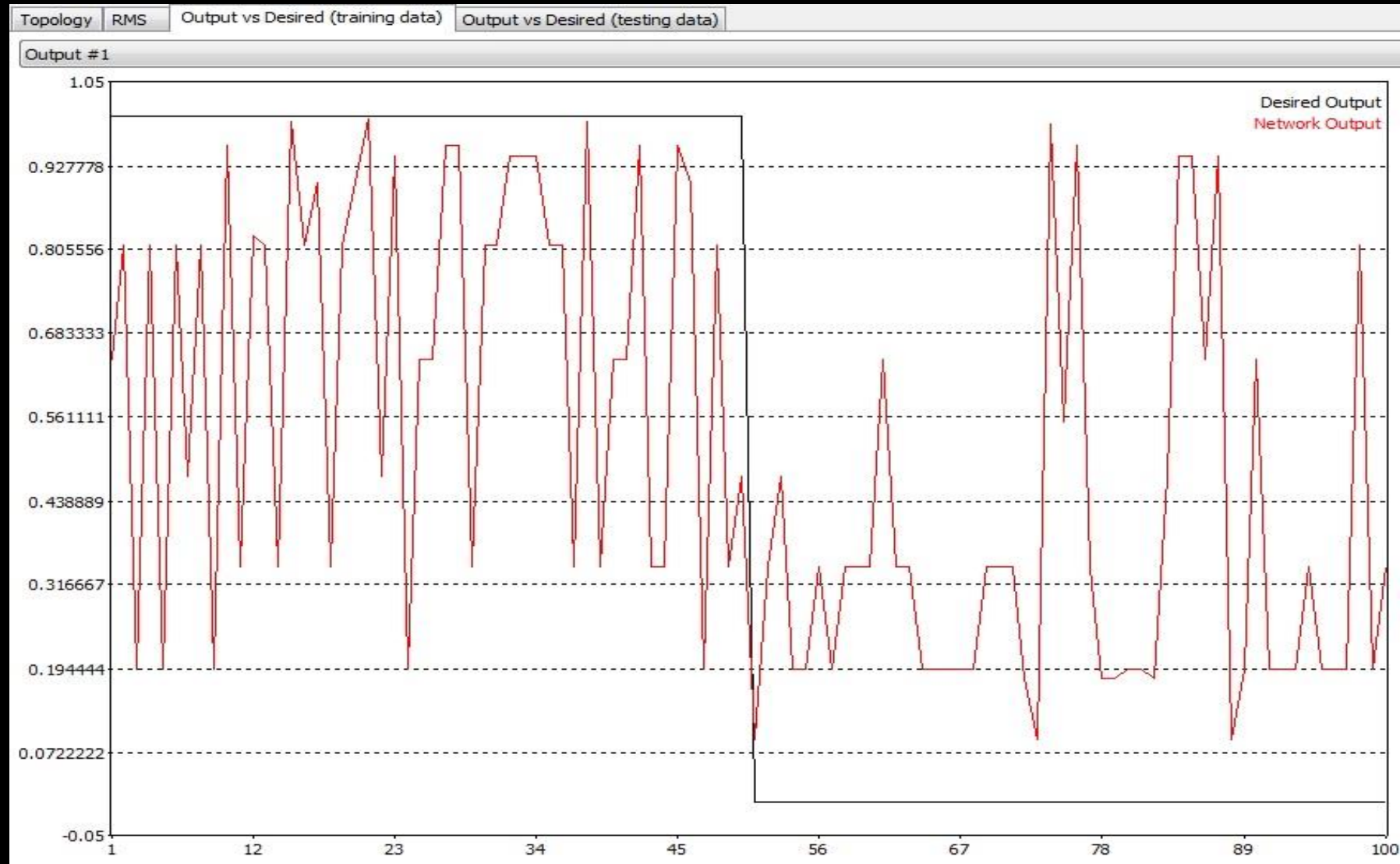


35/63

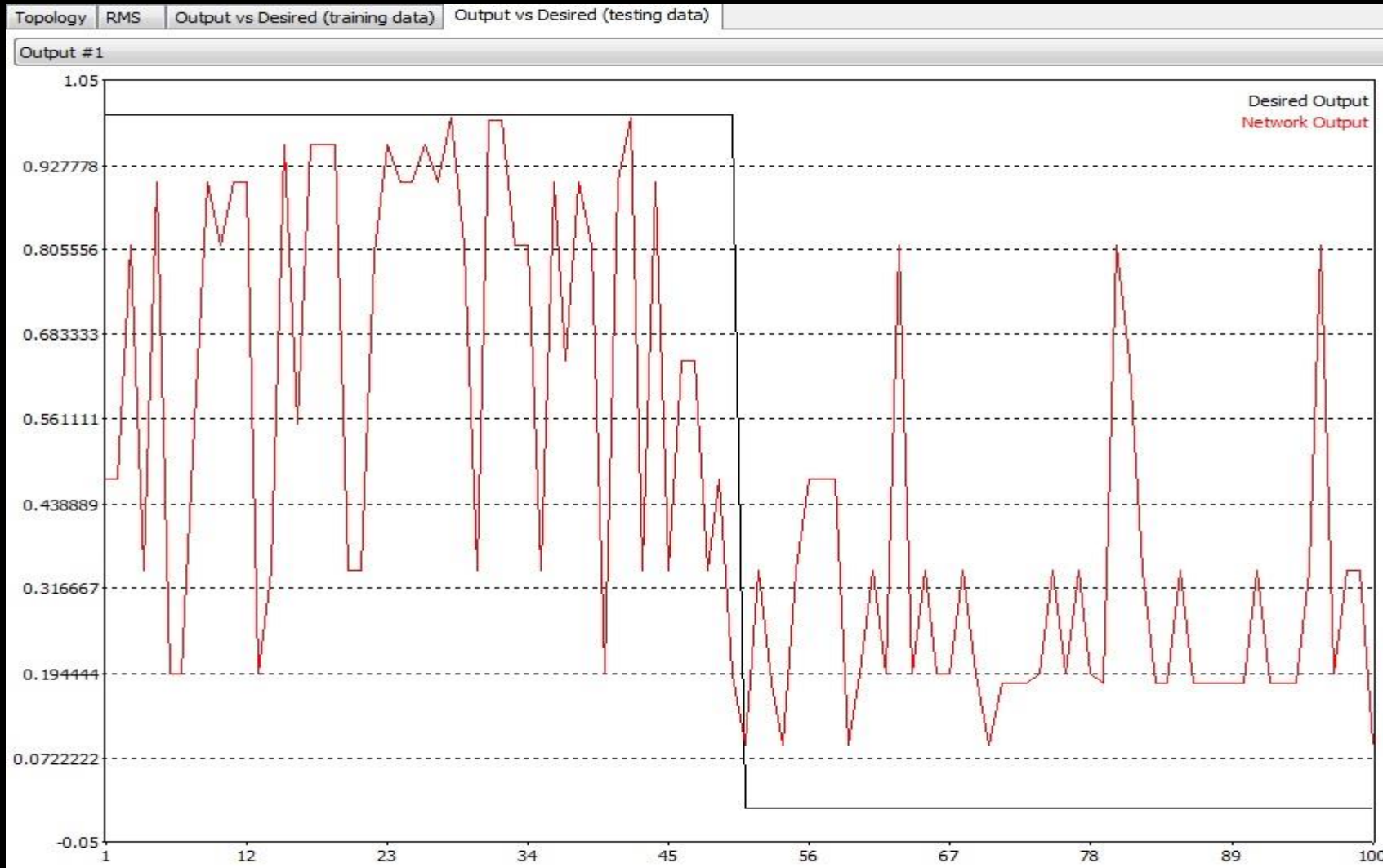
2015

- ▶ RMS Treino: 0,216
- ▶ RMS Teste: 0,191

# Treino: Taxa de acerto 75% (Limiar 0,5)



# Teste: Taxa de acerto 83% (Limiar 0,5)



37/63

2015

# Itens técnicos

- ▶ Coletas
- ▶ Partição de dados de treino e dados de teste
- ▶ Resultados
- ▶ Otimizações da rede
- ▶ Obstáculos

# Estimativa da dose de pigmentos para tintas a partir de uma amostra



# Motivação

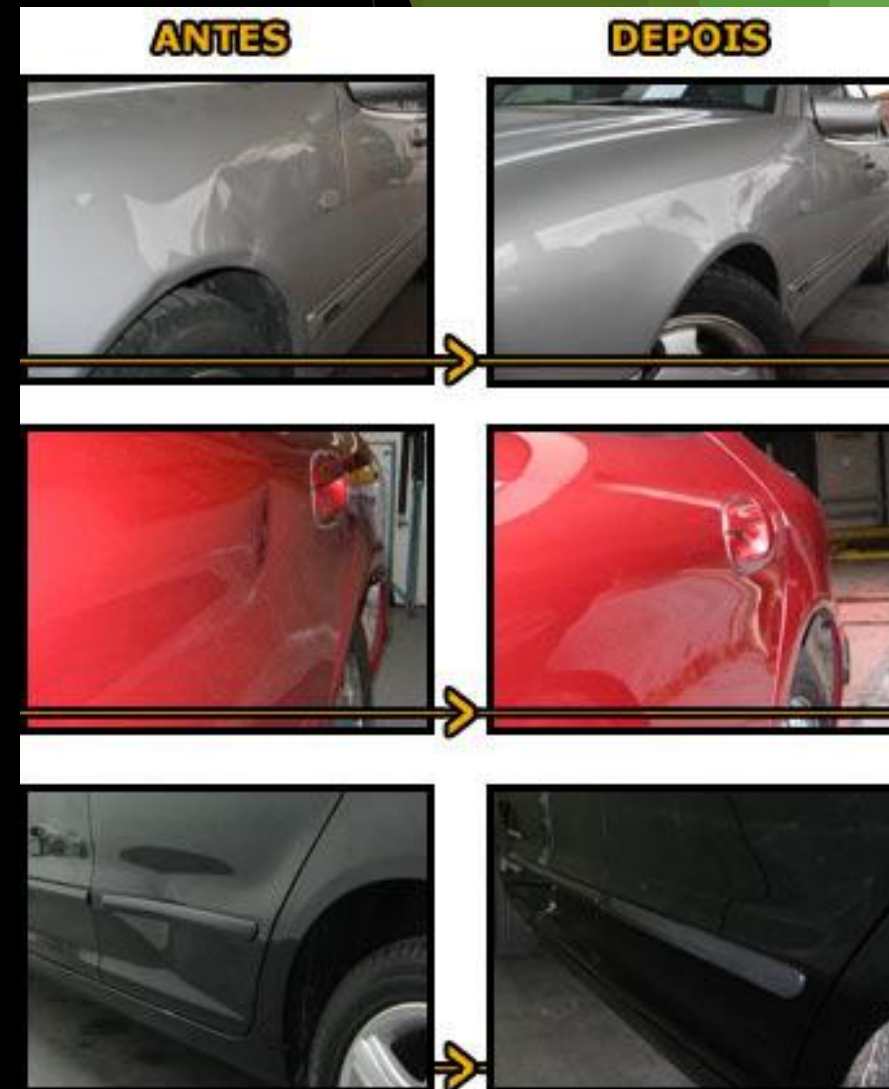
- ▶ como replicar uma determinada cor, dada uma amostra de tinta?
- ▶ Este processo acaba sendo feito por aproximação a olho.





# Exemplo repintura de veículos

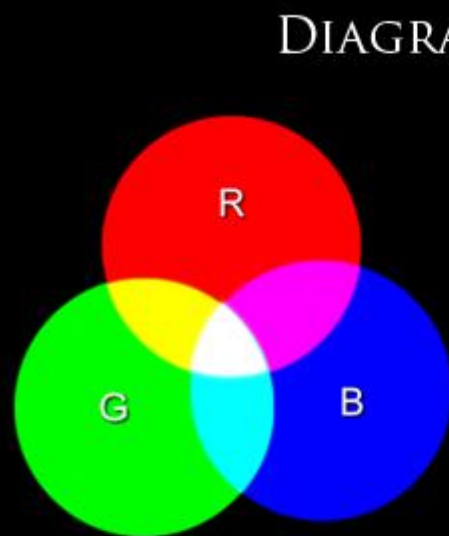
- ▶ Tinta envelhecida
  - ▶ A cor muda com o envelhecimento da pintura
- ▶ Evitar diferenças de tom



- ▶ Tendo em vista essa necessidade, a ideia é criar um dispositivo embarcado portátil que utilize aprendizado de máquina, uma vez que muitas vezes não existe uma fórmula certa para obter determinada cor a partir de um pigmento e existem inúmeros pigmentos de fabricantes diferentes.

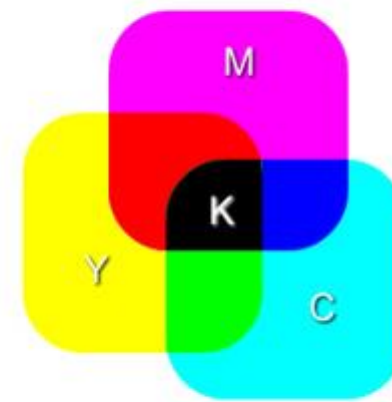
# Sistemas de cor

- ▶ Ser humano possui visão tri-cromática
- ▶ Dois principais sistemas de cor:
- ▶ Aditivo
- ▶ subtrativo



RGB - CORES ADITIVAS

DIAGRAMA DE CORES



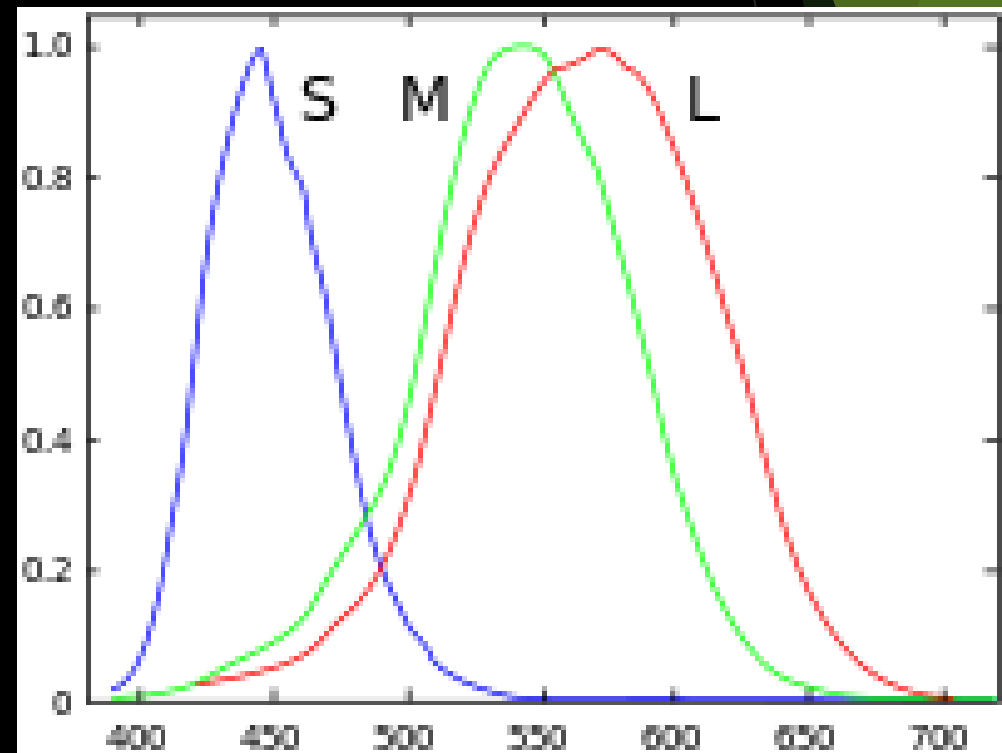
CMYK - CORES SUBTRATIVAS

# Espectro visível de cores

Gama de cores visíveis  
para o olho humano

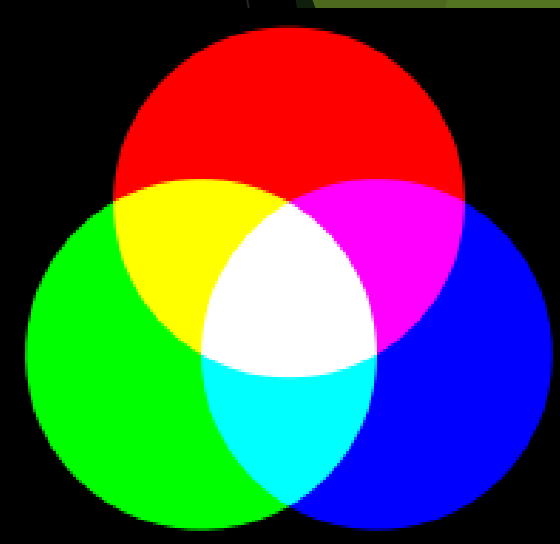


Percepção normalizada  
típica do olho humano



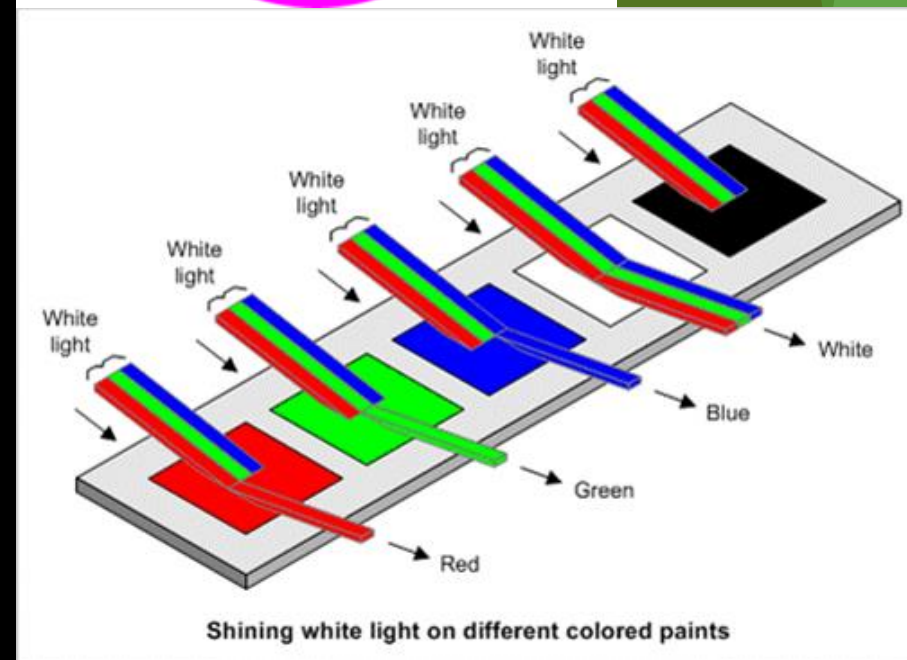
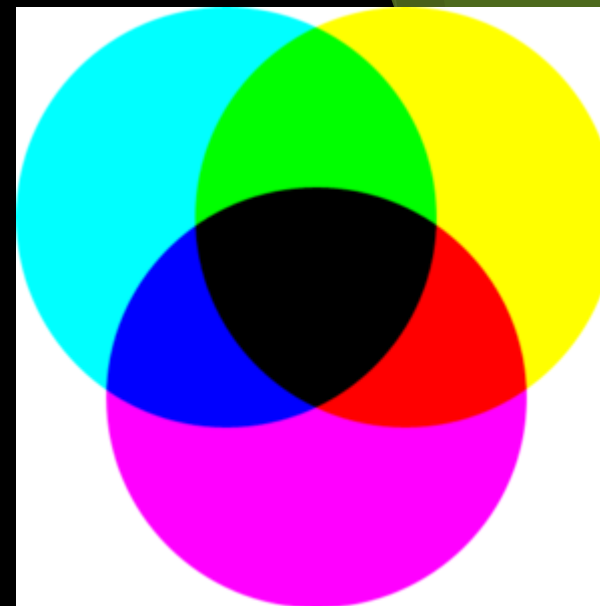
## Sistema aditivo

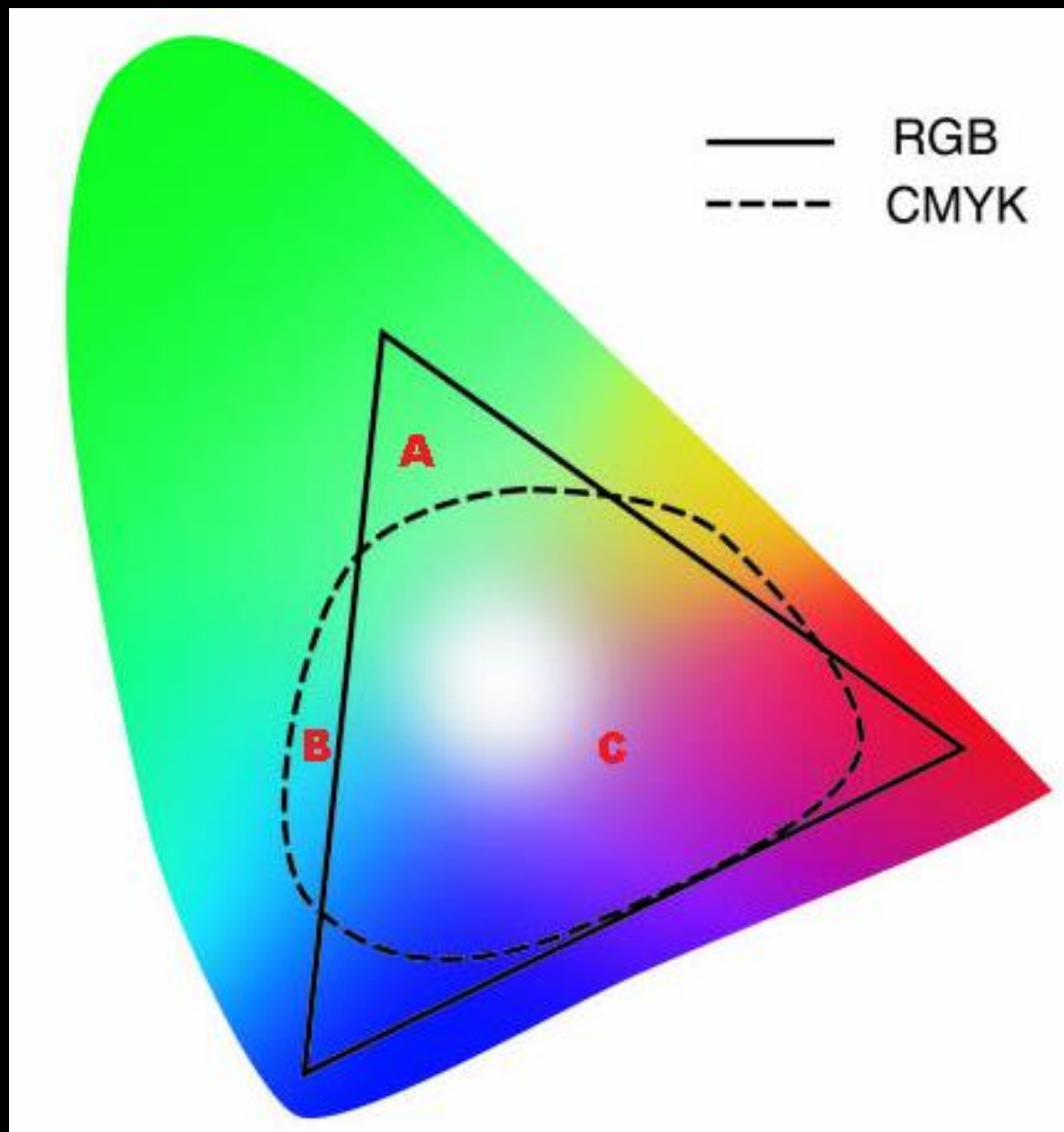
- ▶ Adição das cores para criar novas cores
- ▶ Depende da emissão de luz
- ▶ Principal sistema aditivo é o RGB
- ▶ Reprodução de cores em dispositivos eletrônicos como monitores de TV e computador, "datashows", scanners e câmeras digitais
- ▶ O modelo de cores RGB é baseado na teoria de visão colorida tricromática



# Sistema subtrativo

- ▶ Subtração de cores para criar novas cores
- ▶ Depende da absorção de luz
- ▶ Principal sistema subtrativo é o CMYK
- ▶ Reprodução de cores em mídias impressas e demais sistemas que utilizam pigmentos como jornais revistas e pinturas em geral tingimento de plásticos tecidos estampas.
- ▶ Sistemas comerciais utilizam **Policromia**, utilizam vários pigmentos base
- ▶ Principais sistemas comerciais Pantone, HKS, RAL





# Projeto

- ▶ O sistema faz uso de um sensor de cores RGB que gera um sinal a partir da amostra.
- ▶ O sinal é processado e submetido a uma rede neural artificial em um hardware embarcado
- ▶ A saída da rede determina a quantidade de cada pigmento disponível necessária
- ▶ o resultado é mostrado no LCD.
- ▶ Adaptável para qualquer sistema de cor



Amostra



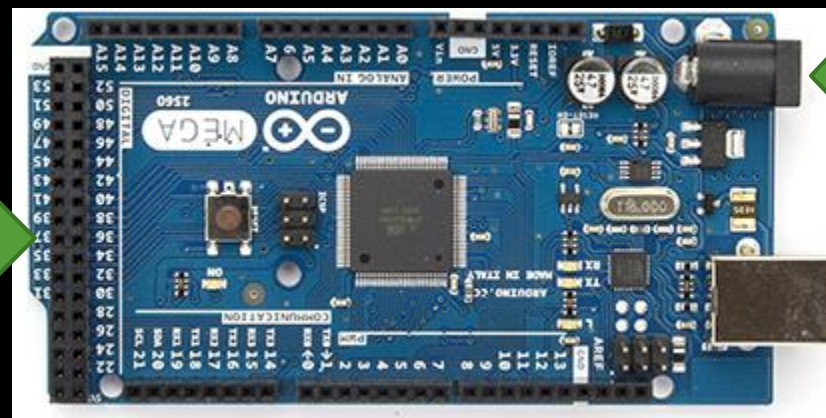
Visor LCD



Alimentação



Sensor



Processador



Baterias

## Processamento do sinal

- ▶ Compensação da resposta dos sensores (função do comprimento de onda).
- ▶ Combinação das diferentes fontes de luz dos sensores para aumentar o número de variáveis de entrada.
- ▶ Remoção de efeitos óticos da superfície do material.
- ▶ Possibilitar a calibração do sensor.

- Exemplo de sensor comercial



## TYPICAL CHARACTERISTICS

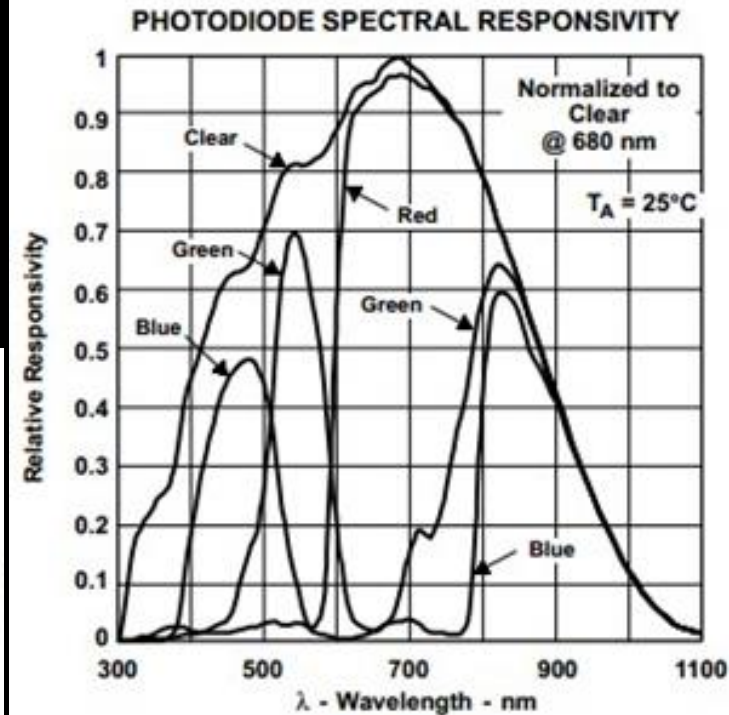


Figure 1

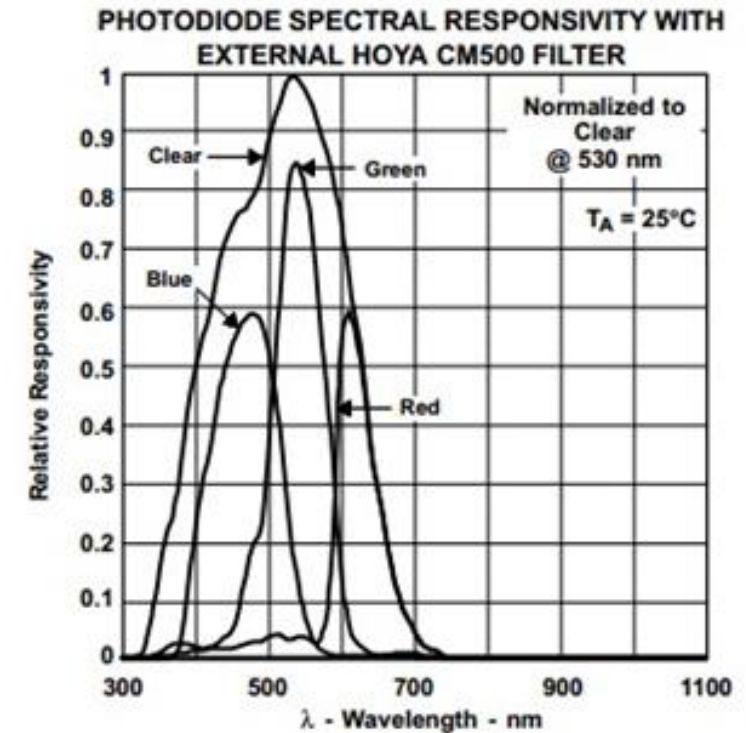


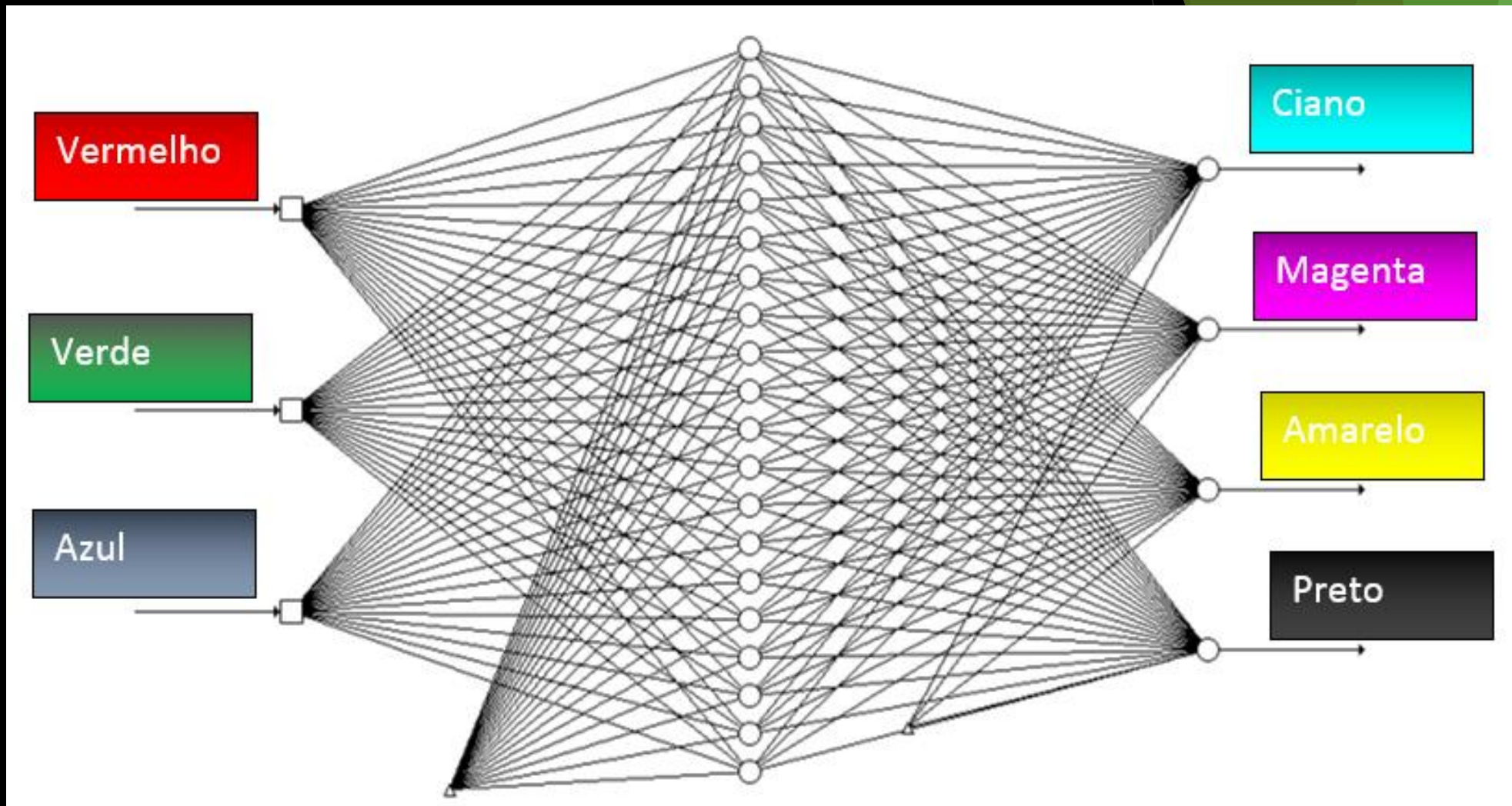
Figure 2

- ▶ Boa relação custo benefício
- ▶ Arduino mega (ou outro arduino)
- ▶ Sensor TC3200
- ▶ Display LCD 16x2
- ▶ Baterias
- ▶ Utilizado CMYK da impressora domestica

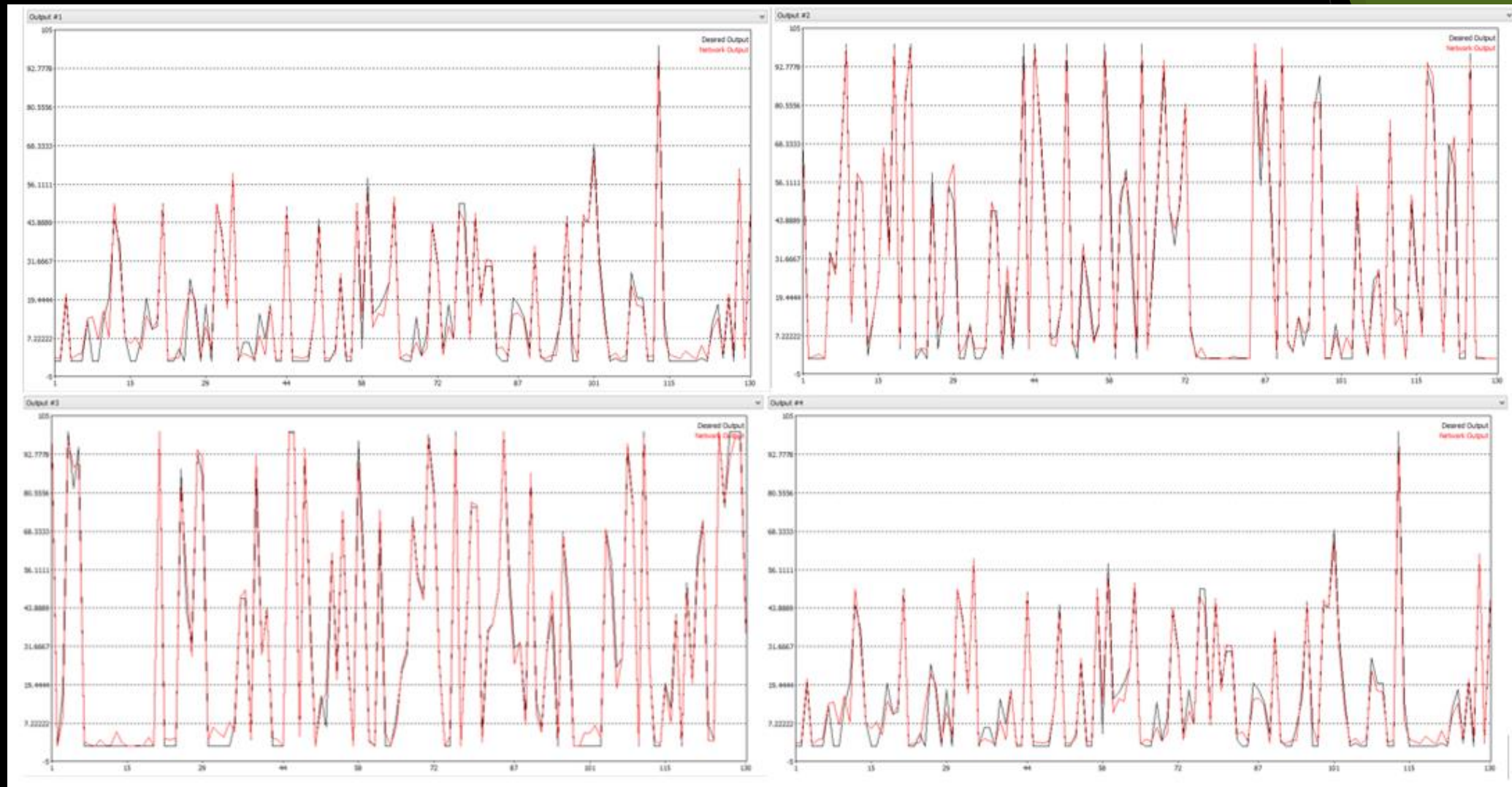


# Detalhes da rede neural

## ► Rede 3-18-4

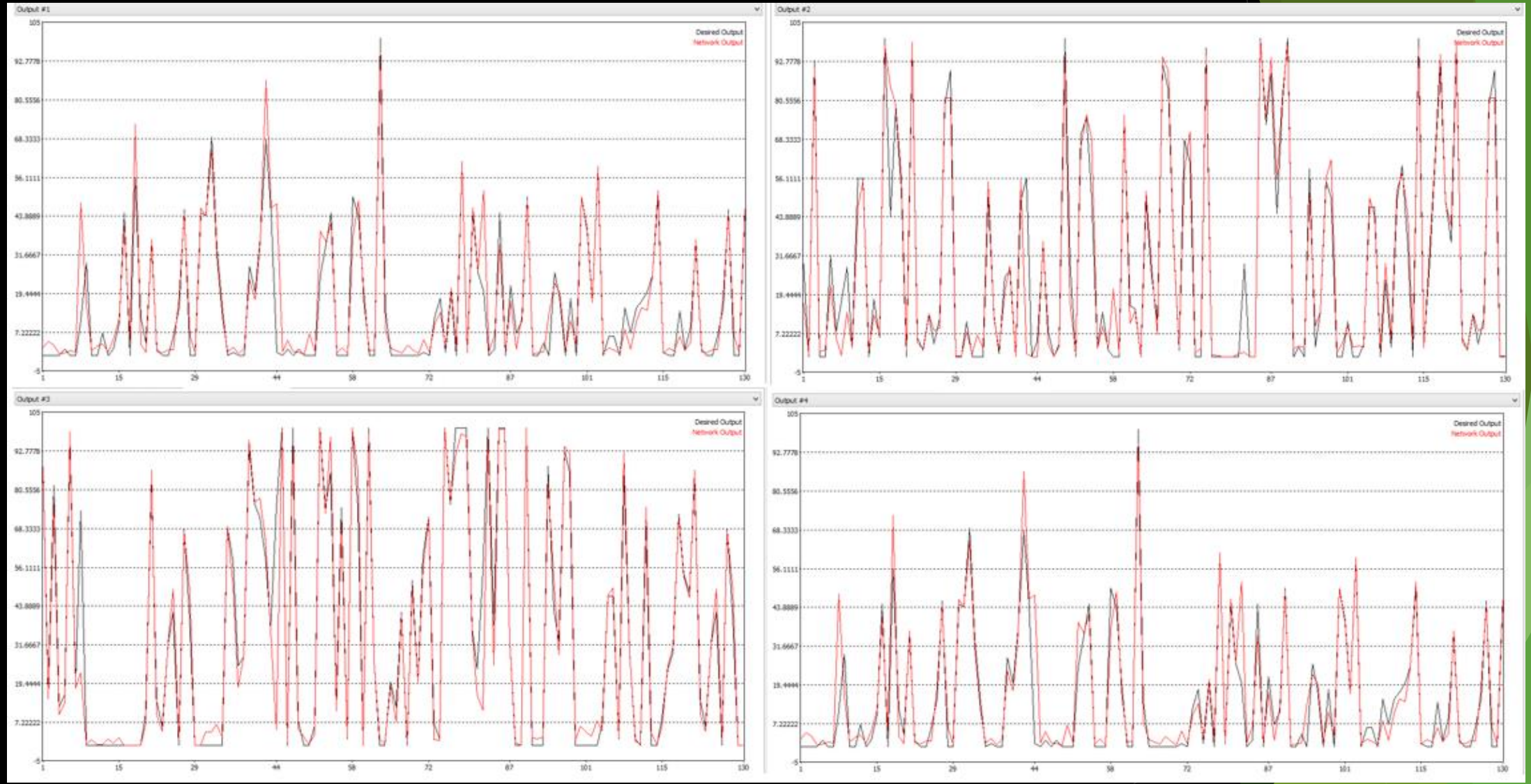


# Treino





# Teste



# Resultados

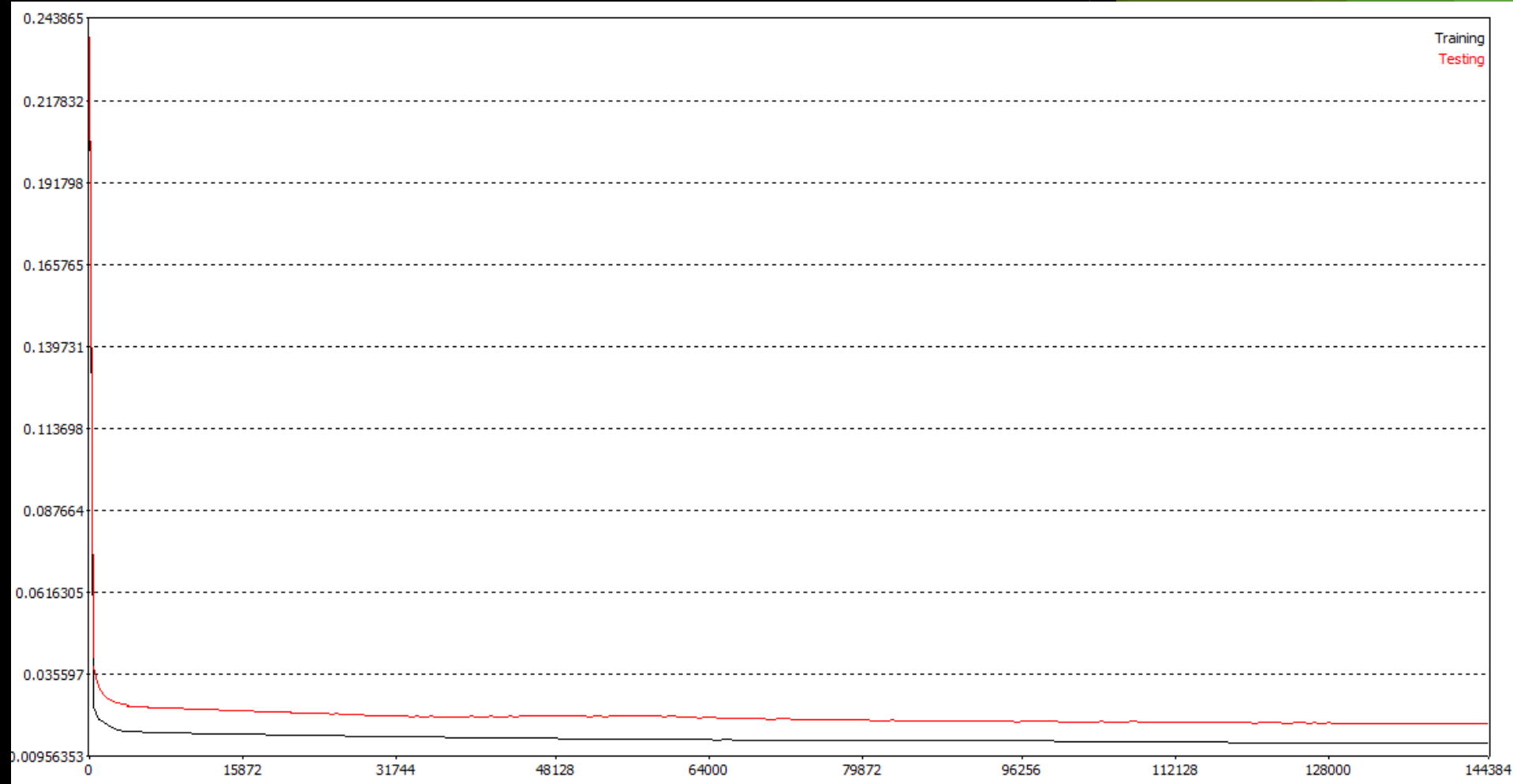
► RMS

► Treino

0,013836873

► Teste

0,017175816



56/63

2015



# Demos

57/63

2015

# Identificação de Fraudes em Compras de Cartão

Programa de avaliação de comportamento de compras no cartão

Nova compra, favor informar:

Valor da compra: 16.36

Data da compra: 25-05-2015

Horário da Compra: 17:59

Categoria da Compra: 5

Forma de Pagamento: 1

Estado da compra: 25

Estado da loja: 25

A rede neural retornou 0.814533, que com limiar de 0.500000 representa uma compra  
a VALIDA

Nova compra, favor informar:

Valor da compra:

# Reconhecimento de Padrões nos Sobreviventes e Vítimas do Titanic

Sexo M/f

M

Idade

0-14 anos? S/N

N

15-35 anos? S/N

S

Passageiro ou Funcionário? P/F

P

passageiro de 1ª? S/N

N

passageiro de 2ª ou 3ª classe

=====

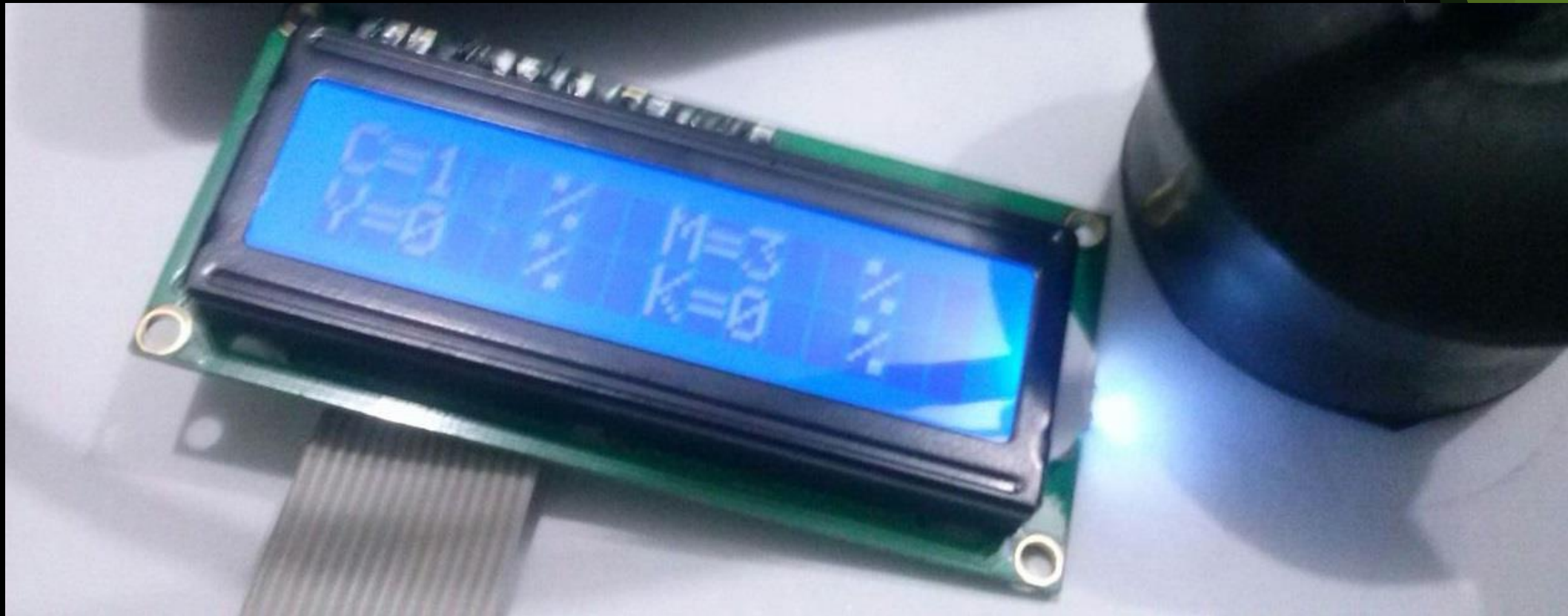
Infelizmente a pessoa não sobreviveria ao naufrágio

Probabilidade de sobrevivência: 24%

Pressione qualquer tecla para continuar. . . \_

# Estimativa da dose de pigmentos para tintas a partir de uma amostra

## ► Amostra de cor branca

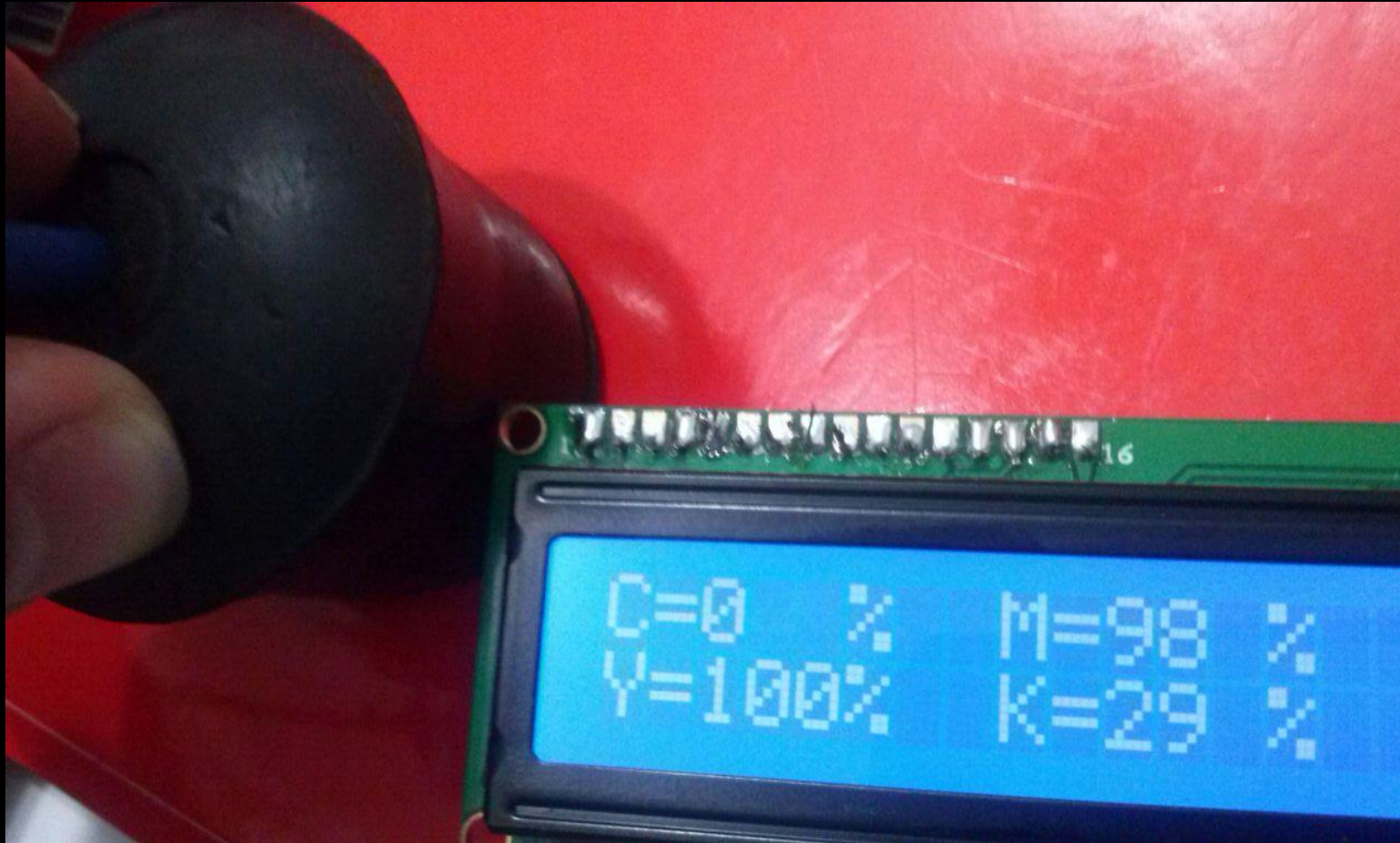


► Amostra de cor preta





► Amostra de cor vermelha



- Amostra de cor amarela

