

SCC0276 - Aprendizado de Máquina

Profa. Dra. Roseli Aparecida Francelin Romero
SCC - ICMC - USP

2019

Sumário

- 1 Introdução
- 2 Avaliação de algoritmos
- 3 Pré-processamento
- 4 Generalização
- 5 Sistema ALVINN

Aplicações de RMC

- Reconhecimento de imagens: faces / gestos;
- Reconhecimento de voz;
- Transformação de texto em voz;
- Previsões da bolsa de valores;
- Análise de dados financeiros;
- Controle automático de robôs.

Tratamento nos dados

- Extração de características;
- Redução de dimensionalidade;
- Construção de características;
- Representação de *sparse*;
- Seleção de características.

Sumário

- 1 Introdução
- 2 Avaliação de algoritmos**
- 3 Pré-processamento
- 4 Generalização
- 5 Sistema ALVINN

Avaliação de algoritmos

- Ciclo básico: dividir o conjunto de dados em treinamento e teste.
- **Holdout:**
$$\begin{cases} p: \text{treinamento} \\ 1 - p: \text{teste} \\ p > 0,5 \end{cases}$$
- **Amostragem aleatória:** L hipóteses são induzidas a partir do conjunto de treinamento.
 - Conjuntos de treinamento e teste: escolhidos aleatoriamente.
 - Erro final: média dos erros de todas as hipóteses induzidas.
Melhor que o holdout.

Avaliação de algoritmos

- **Cross-validation (CV):** os exemplos são divididos em r partições mutuamente exclusivas (*folds*) de tamanho N/r exemplos.
 - Exemplos em $r - 1$ *folds*: treinamento.
 - Exemplos nos *folds* restantes: teste.
 - O processo é repetido r vezes, cada vez com um *fold* diferente para teste.
 - **Erro:** média dos erros calculados em cada um dos r *folds* usados para teste.

Avaliação de algoritmos

- **Stratified cross-validation:** similar ao *cross-validation*, porém mantendo a proporção dos exemplos existentes em cada classe para montar cada *fold*.
- **Leave-one-out:** caso particular do CV, onde 1 exemplo é reservado para teste, e $n - 1$ exemplos, para treinamento.
 - O processo é repetido n vezes.
 - **Erro:** média dos erros em cada teste.

Sumário

- 1 Introdução
- 2 Avaliação de algoritmos
- 3 Pré-processamento**
- 4 Generalização
- 5 Sistema ALVINN

Pré-processamento dos dados

- Uma questão importante em análise de dados.
- Especificar um método: dados sejam descritos de modo que suas características mais importantes sejam ressaltadas.
- Áreas: aprendizado de máquina, processamento de imagens, reconhecimento de voz, estatística.

Pré-processamento dos dados

- Duas metodologias:
 - Métodos estatísticos clássicos:
 - Análise de Componentes Principais (PCA);
 - Análise Discriminante Linear (LDA);
 - Análise de *clusters*.
 - Redes neurais.
- Identificação de faces, compreensão de texto e imagens.

Transformação nos dados

- Dados centrados na média e variância 1 (σ : desvio-padrão):

$$x_m = \frac{1}{n} \sum_{i=1}^n x_i$$
$$x_i(cm) = \frac{x_i - \bar{x}}{\sigma}$$

- Escalamento pela variância (quando temos valor muito dominante em relação aos demais):

$$var_j = \frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2$$
$$x_{ij} \leftarrow \frac{x_{ij}}{var_j}$$

Transformação nos dados

- Pré-processamento dos dados:
 - \log_{10} : enfatiza magnitudes baixas.
 - Normalização:

$$\|x\| = \sum x_{ij}^2 \quad \text{para } i,j=1,\dots,m$$
$$x_{ij}(\text{norm}) = \frac{x_{ij}}{\|x\|}$$

- Não aconselhável para dados com menos de 10 variáveis.
- Multiplicação ou subtração por um número.

Transformação nos dados

- Normalmente, dividem-se todos os elementos do conjunto de dados pelo maior valor, x_{max} .
- Assim, consegue-se que todos os elementos fiquem representados no intervalo $[0, 1]$:

$$x_i \leftarrow \frac{x_i}{x_{max}}$$

Transformação nos dados

- Pode-se dividir todos os dados de uma tabela por:

$$x_{max} - x_{min}$$

- Onde:
 - x_{max} é o **maior** valor do conjunto de dados;
 - x_{min} é o **menor** valor do conjunto de dados.

Sumário

- 1 Introdução
- 2 Avaliação de algoritmos
- 3 Pré-processamento
- 4 Generalização**
- 5 Sistema ALVINN

Generalização

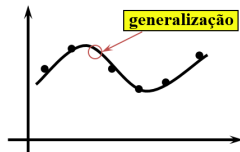
Aprendizado **BP**

conjunto treinamento + algoritmo BP \Rightarrow pesos sinápticos

GENERALIZAR

“GENERALIZAÇÃO”: termo da psicologia.

Processo de aprendizado pode ser visto como um Método de Aproximação de Funções



generalização

: efeito de uma boa aproximação não linear dos dados de entrada, tamanho e eficiência do conjunto treinamento, arquitetura da rede, complexidade física do problema

Complexidade da rede

- **Problema:** determinar o melhor número de nós na camada intermediária.
- Estatisticamente, esse problema é equivalente a determinar o tamanho do conjunto de parâmetros usado para modelar o conjunto de dados. Existe um limite no tamanho da rede.
- Esse limite deve ser tomado lembrando que é melhor treinar a rede para **produzir melhor generalização** do que treinar a rede para representar perfeitamente o conjunto de dados.
- Isso pode ser feito usando **validação cruzada**.

Validação cruzada

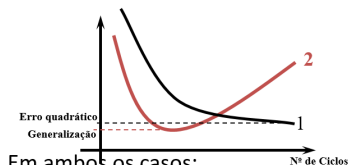
- Conjunto de dados
 - Treinamento ($\sim 75\%$);
 - Teste ($\sim 25\%$).
- Conjunto de treinamento
 - 1 subconjunto: validação do modelo;
 - 1 subconjunto: treinamento.
- Validar o modelo em um conjunto diferente do usado para estimar o modelo.

Validação cruzada

- Usa-se o subconjunto de validação para avaliar o desempenho de diferentes candidatos do modelo (diferentes topologias) e, então, escolher uma delas.
- O modelo escolhido é treinado sobre o conjunto de treinamento inteiro, e a capacidade de generalização é medida no conjunto de teste.

Tamanho do conjunto de treinamento

A validação cruzada pode ser usada para decidir quando o treinamento de uma rede deveria ser cessado.



Curva 1: poucos parâmetros (under fitting)

Curva 2: muitos parâmetros (over fitting)

Em ambos os casos:

- 1) O desempenho do erro na generalização exibe um mínimo
- 2) O mínimo no caso over fitting é menor e mais definido.

Pode-se obter boa generalização se a rede é projetada com muitos neurônios desde que o treinamento é cessado num número de ciclos correspondente ao mínimo da curva do ERRO obtida na Validação Cruzada.

Validação cruzada

- É melhor treinar para produzir a melhor generalização, usando validação cruzada, do que treinar a rede para representar perfeitamente um dado conjunto de dados.

Sumário

- 1 Introdução
- 2 Avaliação de algoritmos
- 3 Pré-processamento
- 4 Generalização
- 5 Sistema ALVINN**

Sistema ALVINN

- Usou uma RN com uma camada intermediária.
 - Entrada: imagem da rodovia (32x30 pixels).
 - Saída: direção.
- **Treinamento:** \sim 10 minutos de direção humana.
- Nenhum controle de velocidade.

ALVINN - Autonomous Land Vehicle in a Neural Network



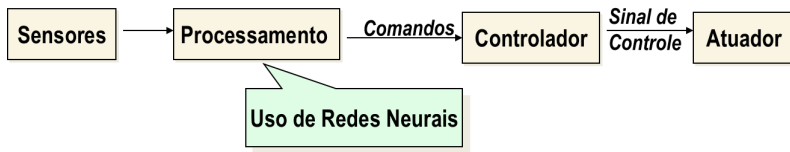
- Usa Rede Neural associada a imagens para guiar um automóvel a altas velocidades (rodovias públicas).

- Dirigiu a velocidades superiores a 100 km/h e por distâncias superiores a 140 km.

- Mostrou a eficiência do uso de uma Rede Neural relativamente simples no controle de um automóvel.

ALVINN - Autonomous Land Vehicle in a Neural Network

- ALVINN é um sistema de percepção para controle de automóveis.
- Demonstrou eficiência do uso de uma rede neural relativamente simples no controle de um automóvel.



ALVINN - Autonomous Land Vehicle in a Neural Network

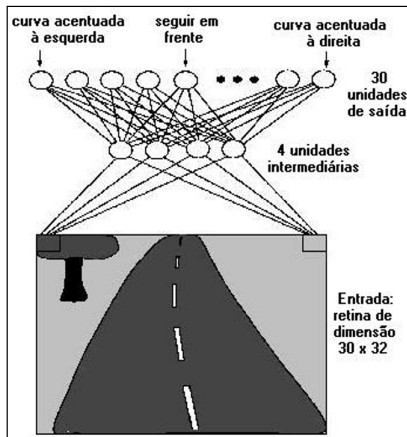
- **Dificuldades:**

- Variações nas formas de pistas (únicas, duplas, não asfaltadas, etc.);
 - Variações na iluminação, devidas às condições de tempo;
 - Processamento deve ser rápido e seguro (tempo real);
 - Requisitos de raciocínio (e.g., para onde virar em junções).
- Redes Neurais Artificiais.
 - Combinação com estruturas simbólicas.

Definição da RNA do ALVINN

- Rede Multi-Layer Peceptron (MLP)
- **Função de ativação:** produz valores entre -1 e 1.
 - **Tangente hiperbólica:** $\varphi(v) = \frac{1 - e^{-v}}{1 + e^{-v}}$
- **Aprendizado:** algoritmo *backpropagation* (paradigma supervisionado).

Definição da RNA do ALVINN



30 neurônios na saída: os mais à esquerda/direita indicam uma virada brusca à esquerda/direita, e os mais ao centro, manter-se em frente.

4 neurônios na camada intermediária: estrutura simples.

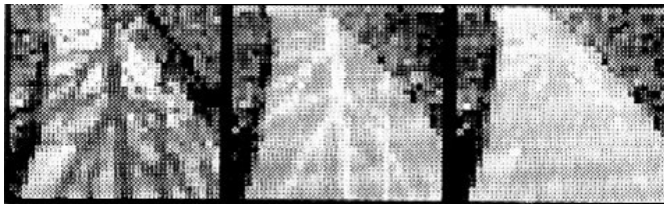
960 neurônios na camada de entrada: como uma "retina" que recebe imagens geradas pelo sensor.

Entradas da RNA

- Alguns processamentos nas informações sensoriais:
 - **Redução** para minimizar gastos computacionais. Capta 480x512 pixels, e faz uma média em subconjuntos (quadros de 16x16).
 - **Normalização (histogram normalization)**: os 5% mais escuros assumem valor -1, enquanto os 5% mais brilhantes, -1, e os demais 90%, valores intermediários proporcionais.

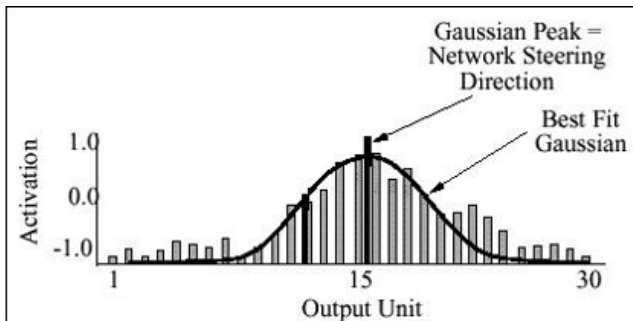
Eliminação de sombras

- Câmera colorida: para **eliminar sombras**, normalizar por brilho em relação à componente azul da imagem.
 - Rodovias possuem, em geral, um alto componente azul.



Saída da RNA

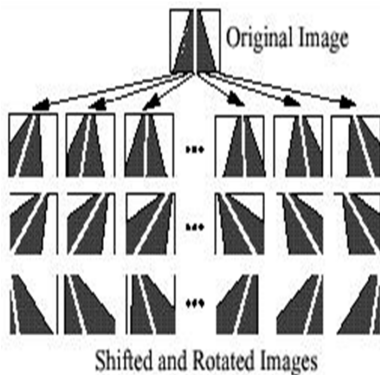
- **Representação gaussiana:** calcula gaussiana que melhor se ajusta à saída (best-fit gaussian).
- Pico da gaussiana indica a direção correta.



Treinamento da RNA

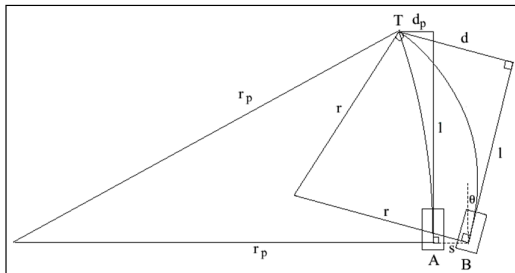
- Muito do sucesso do ALVINN se deve à forma como se deu seu treinamento.
- Treinamento *on-the-fly*: observação de uma pessoa dirigindo.
- Direção correta: a que a pessoa toma.
- **Problemas:**
 - Rede não se defronta com exemplos de comportamentos incorretos e como corrigi-los.
 - Também pode ocorrer *overfitting* (*overlearning*) em relação a padrões recentes.

Treinamento da RNA



- Para aumentar o número de exemplos: transformações nas imagens.
 - Shifts e rotações, de modo a criar 14 novos exemplos.
- Para evitar *overfitting*, ainda usar *buffer*.
- Aproximadamente, quatro minutos de treinamento, a velocidades variando de 8 a 88 km/h.

Obtendo direções



- **T** = posição correta no próximo instante
- **B** = posição transformada em relação a A de s metros e θ graus
- **l** = distância escolhida como ponto final da curva a ser realizada

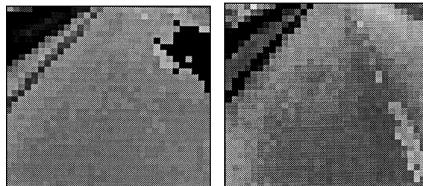
$$r = \frac{l^2 + d^2}{2d}$$

$$d = \cos \theta \cdot (d_p + s + l \tan \theta)$$

$$d_p = r_p - \sqrt{r_p^2 - l^2}$$

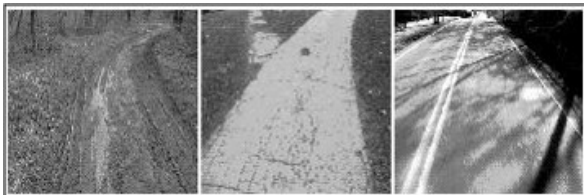
Inserindo ruídos

- Insere ruídos estruturados nas imagens, simulando, por exemplo, carros, *guard-rails*, etc.
- Utiliza conhecimento de que os ruídos aparecem, geralmente, na periferia das imagens.
- Adiciona ruído a 3 de cada 4 representações.



Resultados

- Alguns ambientes:



- Para cada ambiente, uma nova RNA deve ser treinada.
- Flexibilidade de acomodar outros sensores, na impossibilidade de uso da câmera (e.g., à noite).
- Para desvio de obstáculos: usa laser *rangefinder*.