

SCC 5809 REDES NEURAIAS

Profa. ROSELI ROMERO

SCC5809

Tratamento nos dados

- Extração de características
 - redução de dimensionalidade
 - construção de características
 - representação "sparse"
 - seleção de características

Aplicações da RMC

- Reconhecimento de Imagens: FACES/GESTOS
- Reconhecimento de Voz
- Transformação de Texto em Voz
- Previsão de Bolsa de Valores
- Análise de Dados Financeiros
- Controle Automático de Robôs

Avaliação de Algoritmos

- Ciclo Básico
 - Dividir o conj. Dados em $\left. \begin{array}{l} \text{Treinamento} \\ \text{Teste} \end{array} \right\}$
 - **Holdout**: p : treinamento; $(1-p)$: teste; $p > 0.5$
 - **Amostragem Aleatória**: L hipóteses são induzidas a partir do conj. de treinamento
 - conj. trein e teste : escolhidos aleatoriamente
 - erro final: média dos erros de todas as hip. Induzidas
 - melhor que holdout**

- **Cross-Validation (CV)**: os exs. são divididos em r partições mutuamente exclusivas (folds) de tam. N/r exemplos.

Exs. nos $(r-1)$ folds: treinamento

Exs. no fold restante: teste

O processo é repetido r vezes c/ um c/ fold diferente para teste

erro= média dos erros calculados em cada um dos r fold usados para teste

- Diversos sites na internet relacionados com AM

David W.Aha(<http://www.aic.nrl.navy.mil/~aha/research/machine-learning.html>) apresenta varios links para os principais sites da area.

- D.J.Field "What is the goal of sensory coding?" Neural Computation, 6:559-601, 1994
- Mitchell, 1998
- Dietterich, 1997
- Weiss & Kulikowski, 1991

- **Stratified Cross-Validation**: similar ao Cross-Validation porém mantem a proporção dos exs. 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

Pré-processamento dos dados

- Uma questão importante em análise dados
- especificar um método - dados sejam descritos de modo que as suas características mais importantes sejam ressaltadas
- Áreas: Aprendizado de Máquina, Processamento de Imagens, Reconhecimento de voz, estatística.

Pré-processamento dos Dados

- Duas metodologias:
 - métodos clássicos estatísticos:
 - Análise de Componentes Principais (PCA)
 - Análise Discriminante Linear (LDA)
 - Análise de Clusters
 - Redes Neurais
- identificação de faces, compressão de texto e imagens.

Transformação nos dados

- Pré-processamento dos dados
 - log 10 (enfatiza magnitudes baixas)
 - Normalização
 - $||x|| = \sqrt{\sum x_{ij}^2}$ para $i, j = 1, \dots, m$
 - $x_{ij(\text{norm})} = x_{ij} / ||x||$
 - não aconselhável para dados com menos de 10 variáveis
 - multiplicação/ subtração por um no.

Transformação dos dados

- Dados centrados na média e variância 1
 - $x_m = 1/n \sum_{i=1}^n x_i$
 - $x_i(\text{cm}) = (x_i - x_m) / \sigma$ (σ : desvio padrão)
- Escalamto pela variância (quando temos muito dominante em rel. as demais)
 - $\text{var}_j = 1 / (n-1) \sum_{i=1}^n (x_{ij} - x_{mj})^2$
 - $x_{ij} = x_{ij} / \text{var}_j$

Transformação nos dados

- Normalmente se divide todos os elementos do conj. de dados pelo maior valor, x_{max} . Assim, se consegue que todos os elementos fiquem representados no intervalo [0,1]:
 $x_i = x_i / x_{\text{max}}$

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

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 a melhor generalização** do que treinar a rede para representar perfeitamente um conjunto de dados.

Isso pode ser feito usando validação cruzada.

Generalização

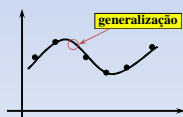
Aprendizado BP

conjunto treinamento + algoritmo BP \Rightarrow pesos sinápticos

GENERALIZAR

"GENERALIZACAO": 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

Validação Cruzada

- Conjunto de dados:
 - treinamento (~75%)
 - teste (~25%)
 - Conjunto de Treinamento
 - 1 subconjunto: validação do modelo
 - 1 subconjunto: treinamento
- Validar o modelo num conj. diferente do usado para estimar o modelo

Validação Cruzada

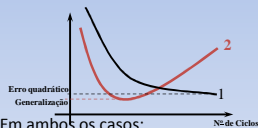
- Usa-se o subconjunto de validação para avaliar o desempenho de diferentes candidatos do modelo (dif. Topologias) e então escolhe-se uma delas.
- O modelo escolhido é treinado sobre o conj. treinamento inteiro e a capacidade de generalização é medida no conjunto de teste.

Validação Cruzada

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

Tamanho do Conjunto 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.

- [Sistema ALVINN](#)

Usou uma RN com uma camada intermed.

-INPUT: Imagem da rodovia (32 x 30 pixels)

-SAÍDA: Direção

-Treinamento: ~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

• Dificuldades:

- variações nas formas de pistas (únicas, duplas, não asfaltadas, etc);
- variações na iluminação e condições de tempo;
- processo deve ser rápido e seguro (tempo real);
- requisitos de raciocínio (ex.: para onde virar em junções)

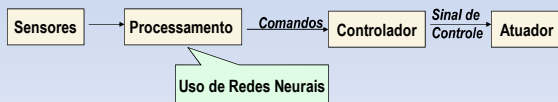
Redes Neurais Artificiais

Combinação com estruturas simbólicas

ALVINN

Autonomous Land Vehicle in a Neural Network

- ALVINN é um sistema de percepção para controle de automóveis
- Demonstrou a eficácia do uso de uma Rede Neural relativamente simples no controle de um automóvel

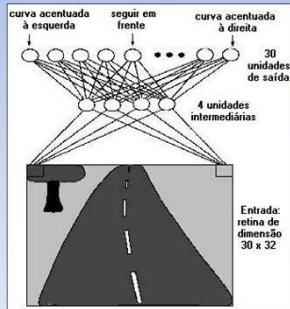


Definição da RNA do ALVINN

- Rede Multilayer Perceptrons (MLP)
- **Função de Ativação:** produz valores entre -1 e 1
Tangente Hiperbólica - $\phi(v) = \frac{1 - \exp(-v)}{1 + \exp(-v)}$
- **Aprendizado:** Algoritmo back-propagation (Paradigma Supervisionado).

Definição da RNA do ALVINN

Arquitetura



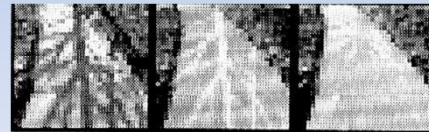
30 neurônios na saída: os mais à esquerda/direita indicam uma virada brusca à esquerda/direita, e o mais ao centro, manter-se em frente.

4 neurônios na camada intermediária: estrutura simples.

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

Eliminação de sombras

– *Câmera colorida:* para **eliminar sombras** normaliza por brilho em relação à componente Azul da imagem (rodovias possuem em geral um alto componente azul).



Entradas da RNA

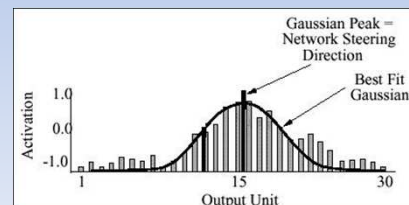
- **Alguns processamentos nas informações sensoriais:**

- **Redução** para minimizar gastos computacionais. Capta em 480x512 pixels, e faz uma média em subconjuntos (quadros de 16x16).

- **Normalização (Histogram Normalization):** 5% mais escuros assumem valor -1 , 5% mais brilhantes, $+1$ e os demais 90%, valores intermediários proporcionais.

Saída da RNA

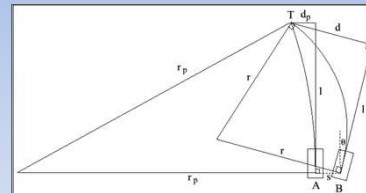
- **Representação Gaussiana:** calcula Gaussiana que melhor se ajusta à saída (**best-fit Gaussian**). Pico da Gaussiana indica a direção correta.



Treinamento da RNA

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

Obtendo direções



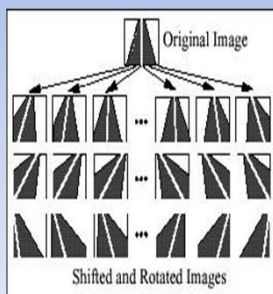
$$r = \frac{l^2 + d^2}{2d}$$

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

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

- A = posição correta
- 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

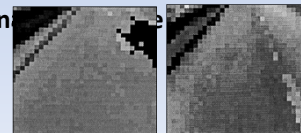
Treinamento da Rede Neural



- Para aumentar o conjunto de exemplos \Rightarrow transformações nas imagens.
- Shifts e Rotações de forma a criar 14 novos exemplos.
- Para evitar overfitting, ainda usa buffer.
- Aproximadamente quatro minutos de treinamento a velocidades variando de 8 a 88 km/h.

Inserindo ruídos

- Inserir Ruídos Estruturados nas imagens, simulando, por exemplo, carros, guardrails, etc.
- Utilizado conhecimento de que os ruídos aparecem geralmente na periferia das imagens.
- Adiciona transformações.



Resultados

- Alguns ambientes:



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