

-
-
-
-
-

IN0997 - Redes Neurais Fundamentos

Aluizio Fausto Ribeiro Araújo
Universidade Federal de Pernambuco
Centro de Informática - CIn
Departamento de Sistemas da Computação
aluizioa@cin.ufpe.br



Introdução

- Este tópico visa apresentar sucintamente alguns princípios neurocognitivos que servem de base para proposição dos modelos de redes neurais artificiais (RNA). Tais modelos utilizam abstrações de neurônios reais que são as unidades básicas responsáveis pelo processamento cerebral.
- Este tópico visa também apresentar um conjunto de requisitos compartilhadas por um grande número de modelos de redes neurais artificiais. Os requisitos são organizados em um arcabouço (*framework*) comum a maioria dos modelos de redes neurais artificiais.
- O último objetivo deste tópico é discutir uma metodologia para utilização de redes neurais em problemas do mundo real.

Conteúdo

- Princípios Neurocognitivos
 - Informações sobre neurônios;
 - Neurônio típico;
 - Comportamento elétrico neuronal.
- Arcabouço geral
 - Definição do arcabouço;
 - Detalhes de cada componente.
- Metodologia para utilização de uma rede neural
 - Definição dos passos da metodologia;
 - Detalhes de cada passo.

Princípios neurocognitivos: Motivação

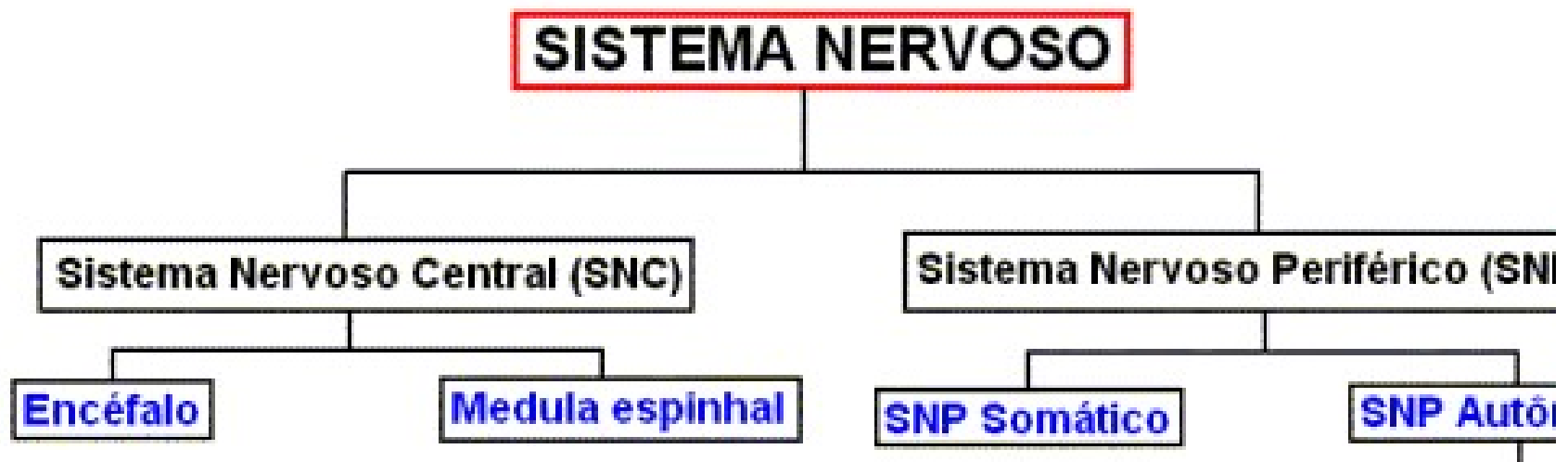
- Para se propor modelos de redes neurais artificiais (RNA), em geral, precisa-se conhecimento em:
 - Neurociência: Hipóteses para serem utilizadas nos modelos.
 - Ferramentas para Modelagem: Representam, descrevem, analisam e projetam os modelo para as hipóteses.
 - Ciência cognitiva provê comportamentos que o sistema deve ter.
- Este tópico visa apresentar sucintamente alguns princípios neurocognitivos que fundamentam proposição das RNAs.
 - Tais modelos utilizam abstrações de neurônios reais que são as unidades básicas responsáveis pelo processamento cerebral.

Princípios neurocognitivos: Motivação

- Características do cérebro que são desejáveis em sistemas artificiais:
 - Robustez e tolerância à falhas.
 - Degradação paulatina sob danos.
 - Flexibilidade (capacidade de se adaptar a novos ambientes).
 - Habilidade para lidar com informação difusa, probabilística, ruidosa e inconsistente.
 - Alto grau de paralelismo.
 - Tamanho compacto.
 - Alto aproveitamento energético.

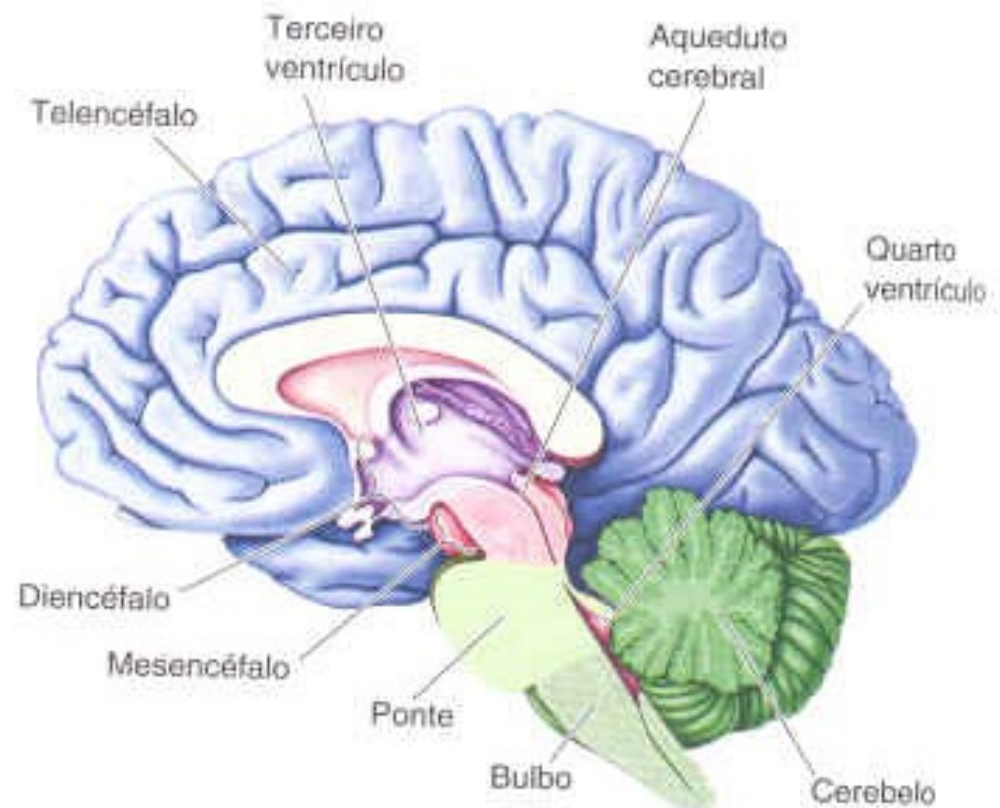
Princípios neurocognitivos: Motivação

- O Sistema Nervoso Central (SNC) recebe, analisa e integra informações. É o local onde ocorre a tomada de decisões e o envio de ordens. O Sistema Nervoso Periférico (SNP) carrega informações dos órgãos sensoriais para o sistema nervoso central e do sistema nervoso central para os órgãos efetores (músculos e glândulas).



Princípios neurocognitivos: Motivação

- O SNC é dividido em encéfalo e medula. O encéfalo é formado pelo telencéfalo (hemisférios cerebrais), diencéfalo (tálamo e hipotálamo), cerebelo, e tronco cefálico. Este último é composto por bulbo, mesencéfalo e ponte.



Princípios neurocognitivos: Motivação

- Bilhões de corpos celulares de neurônios estão na região superficial do telencéfalo (substância cinzenta). Tal região é denominada córtex cerebral, formado a partir da fusão das partes superficiais telencefálicas e diencefálicas.



- O córtex está envolvido em controle do movimento, memória e função cognitiva.

Informações sobre neurônios

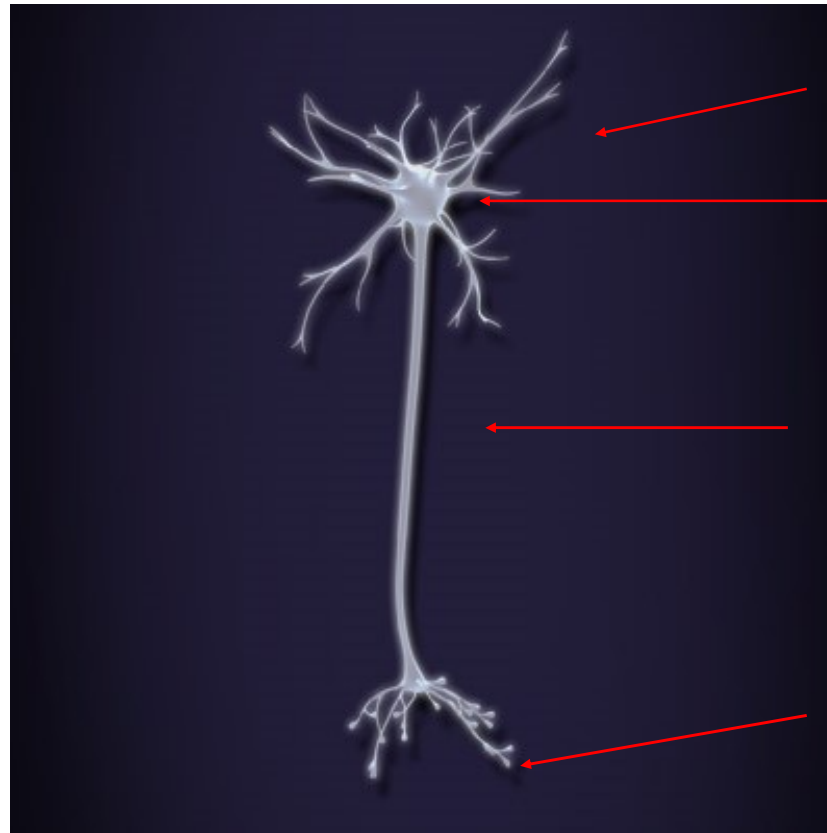
- Os vários tipos de neurônios se caracterizam por:
 - Possuírem os mesmos genes, a mesma organização geral e o mesmo aparato bioquímico.
 - Diferirem basicamente na forma da célula, na membrana externa capaz de gerar impulsos nervosos, na presença de uma estrutura única para transferir informações entre neurônios.
 - Deste modo, não existe dois neurônios iguais.
- Um ser humano adulto:
 - Possui entre 10^{10} e 10^{11} neurônios.
 - Cada neurônio possui centenas ou milhares de conexões e estão protegidos de estimulação mecânica pelo crânio.

Informações sobre neurônios

- Os neurônios são metabolicamente muito ativos. O sistema nervoso central consome cerca de 25% da energia do corpo embora pese de 1% a 2% de seu peso. Além disto, os neurônios são também sensíveis a intoxicação e falta temporária de combustível e são protegidos por mecanismos de filtragem.
- Estas células não se dividem logo após o nascimento. Antes dele, ocorre morte celular programada que é aparentemente causada pela não inserção de neurônios em redes. Além disto, a evolução procura manter o menor número possível de neurônios pois sua operação custa caro.

O neurônio típico

- Um neurônio é uma célula composta de um corpo celular (onde estão núcleo, citoplasma e citoesqueleto) e de seus prolongamentos celulares finos, os neuritos, que podem ser subdivididos em dendritos e axônios.



Dendritos

Corpo
celular

Axônio

Sinapses

- A região de passagem do impulso nervoso de um neurônio para a célula adjacente é chamada de sinapse.

O neurônio típico

- O terminal de entrada dos neurônios têm processadores finos: os dendritos. Estes são estruturas tubulares que se ramificam para formar um ramo de uma árvore ao redor do corpo celular. Os dendritos constituem a superfície física através da qual um neurônio recebe sinais de entrada de muitas outras células.
- O corpo celular ou soma contem o núcleo da célula e os mecanismos para executar a síntese de moléculas, essencial a sobrevivência celular. O corpo celular é caracterizado por: não se dividir depois da fase embrionária; projetar vários dendritos e um único axônio a partir de sua estrutura; ser recoberto por sinapses juntamente com os dendritos; possui mecanismos para prover energia à célula e sintetizar proteínas.

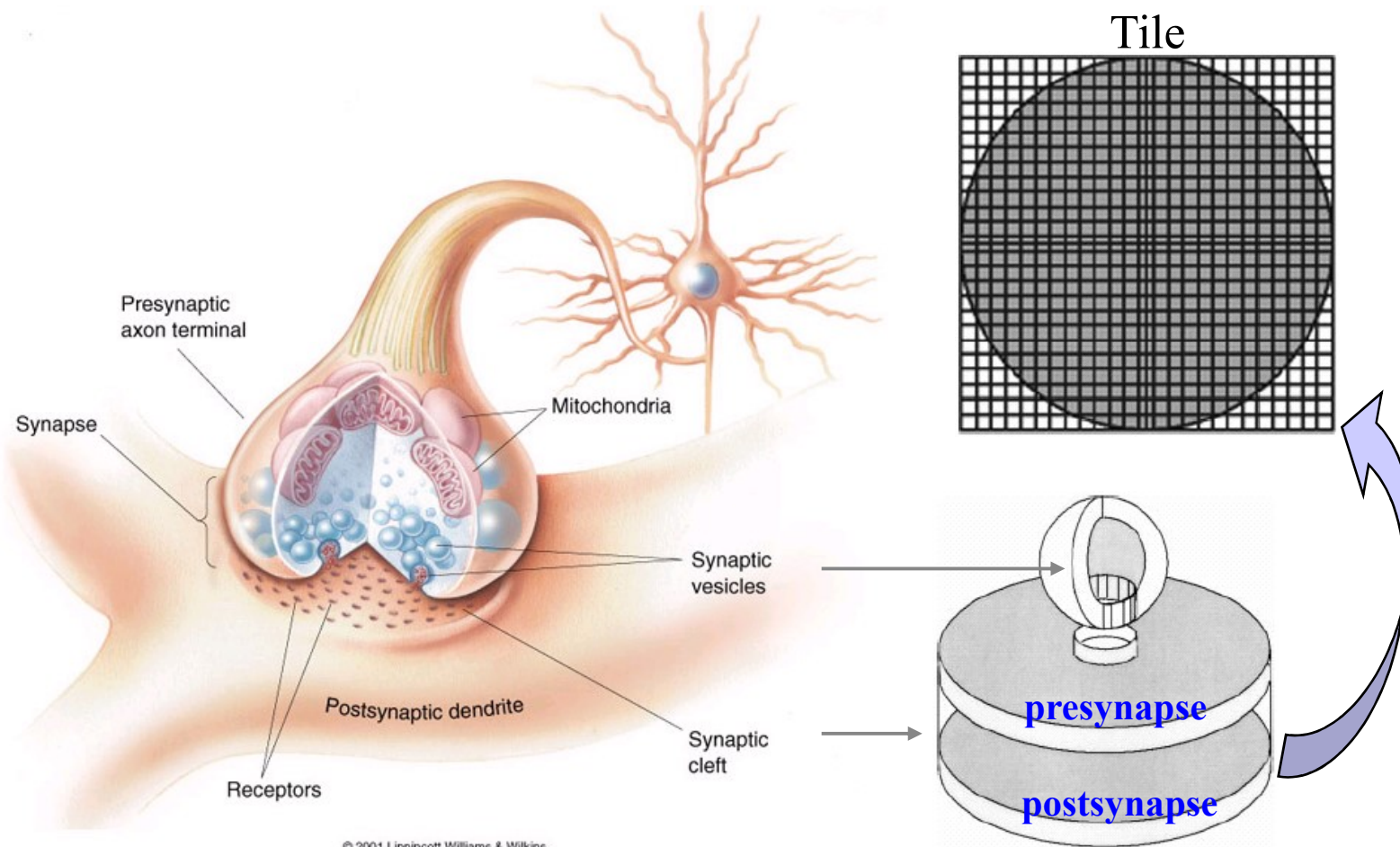
O neurônio típico

- O axônio é a “linha de transmissão” do neurônio. Eles são extensões do corpo celular nos quais sinais trafegam a partir deste corpo por longas distâncias. Em geral os axônios se ramificam na parte final (arborização terminal) permitindo que a informação de uma célula atinja muitas outras. O axônio difere dos dendritos estruturalmente e nas propriedades de sua membrana externa. Os axônios são mais finos e longos que os dendritos e ramificam-se apenas na parte terminal. A membrana externa dos axônios propaga impulsos elétricos e libera neurotransmissores em seus terminais enquanto que a membrana nos dendritos responde aos neurotransmissores.

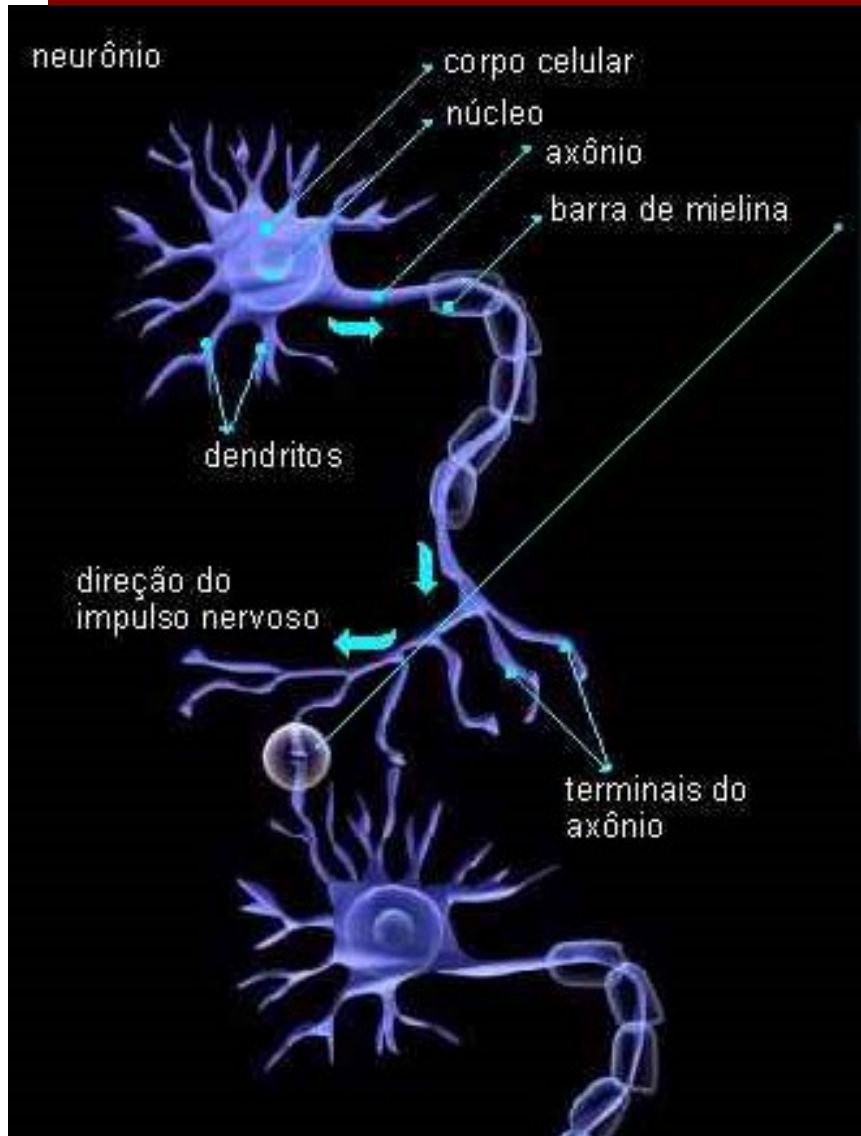
O neurônio típico

- As sinapses são pontos de contato especializados nos quais a informação é transferida de uma célula para outra. As sinapses ocorrem entre axônio e dendrito (maioria), axônio e axônio, dendrito e dendrito e axônio e corpo celular. As sinapses permitem uma célula influenciar uma outra em força e natureza. O axônio aferente é chamado de célula pré-sináptica enquanto que o dendrito eferente é chamado de célula pós-sináptica.

O neurônio típico - sinapse



O neurônio típico

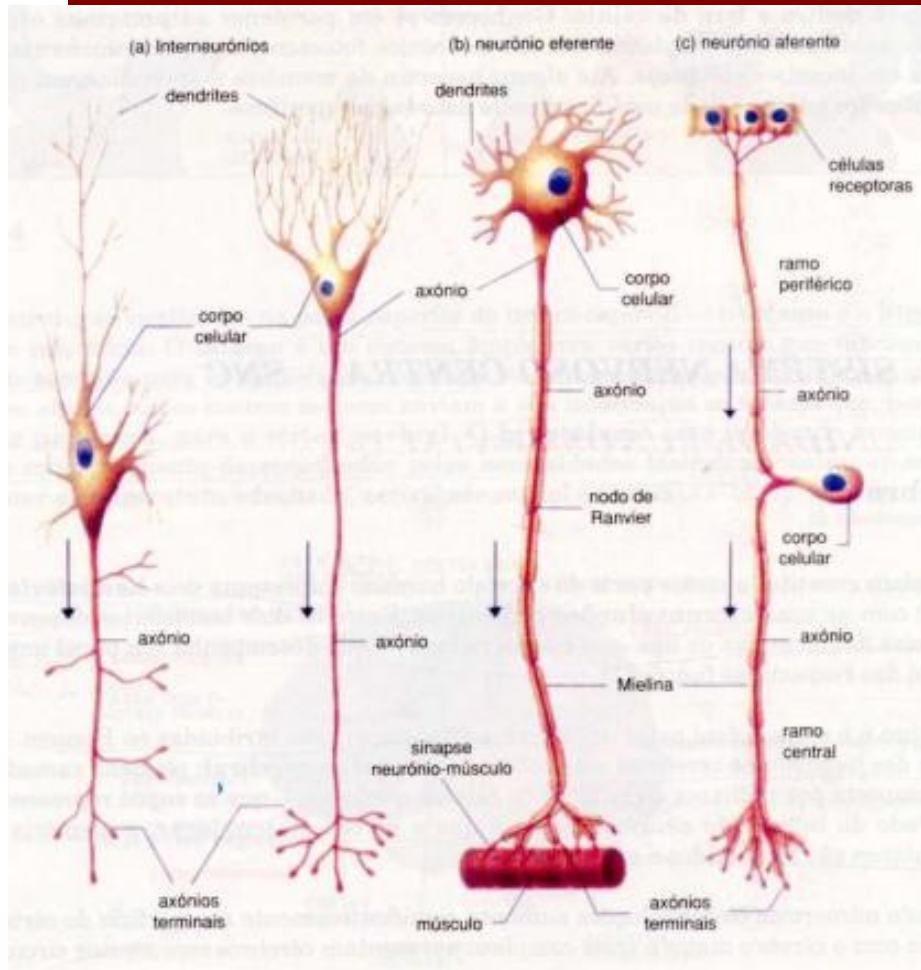


- Um conjunto de sinais vindo dos dendritos são processados no corpo celular de um neurônio. Se ele atingir uma certa diferença de potencial entre o meio interno e o meio externo, um impulso (ou um trem deles) é propagado pelo axônio. Nas extremidades de axônio, o sinal é transmitido para os dendritos de um ou mais neurônios. Neste ponto, recomeça, para o novo neurônio, o processamento explicado.

O neurônio típico

- Os neurônios se diferenciam por:
 - Tipos de neurotransmissores
 - Estruturas dendríticas;
 - Axônios recobertos por *myelin*;
 - Tamanhos diferentes;
 - Tamanhos de razões (e.g., alguns com axônios muito longos, até 1m).

O neurônio típico



- Classificação de neurônios de acordo com suas funções na condução dos impulsos:
- 1. **Neurônios receptores ou sensitivos (aferentes):** Recebem estímulos sensoriais e conduzem o impulso nervoso ao sistema nervoso central.
- 2. **Neurônios motores ou efetadores (eferentes):** Transmitem os impulsos motores (respostas ao estímulo).
- 3. **Neurônios associativos ou interneurônios:** Estabelecem ligações entre neurônios receptores e motores.

Comportamento Elétrico Neuronal

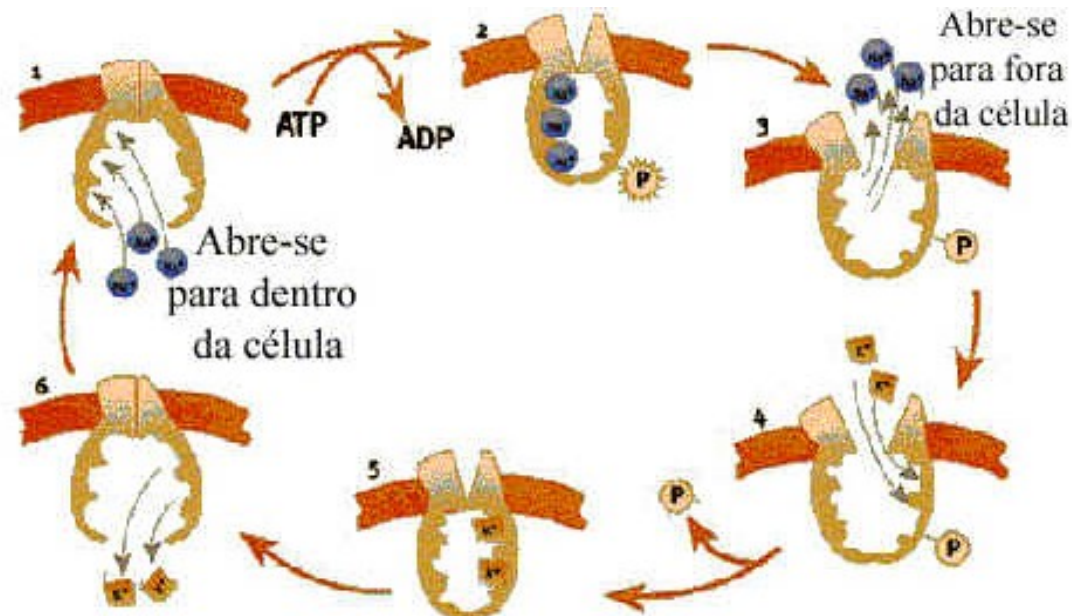
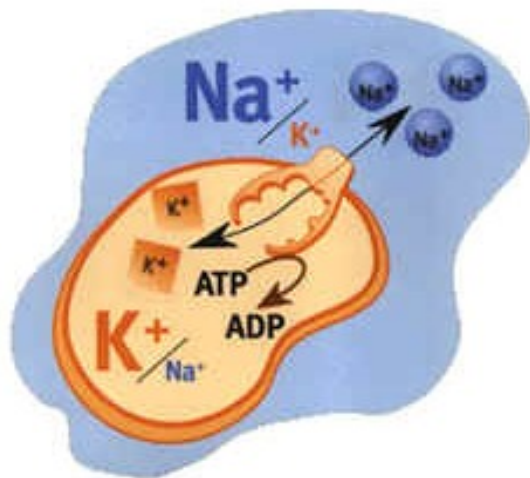
- A membrana celular separa a parte externa da célula da interna. Ela mede entre 60 e 70 Å e é basicamente composta de lipídeos e proteínas. Proteínas passam através da membrana via poros que se modificam na presença de excitações elétricas ou químicas. A informação flui através das sinapses e envolvem neurotransmissores.
- As partes interna e externa da membrana possuem uma diferença de potencial elétrico, tipicamente situado entre -50 e -90 mV. A parte interna é mais negativa que a externa. Esta diferença de potencial é causada por concentrações de eletrodos distintas em cada meio.

Comportamento Elétrico: Concentrações Iônicas

- As concentrações de sódio e potássio são diferentes dentro e fora da membrana. A quantidade de sódio é 10 vezes menor dentro que fora da membrana, enquanto que a quantidade de potássio no lado interno à membrana é 40 vezes maior que aquela do lado externo. Estas concentrações são mantidas por uma bomba de sódio, mecanismo que se situa nas membranas. Esta bomba move sódio para fora e potássio para dentro da membrana por um complexo de moléculas de proteínas grandes, as custas de energia metabólica. A bomba funciona ininterruptamente. Esta situação gera um estado de repouso da célula, no qual a diferença de potencial compensa a diferença iônica.

Comportamento Elétrico: Concentrações Iônicas

- A membrana plasmática do neurônio transporta alguns íons ativamente, do líquido extracelular para o interior da fibra, e outros, do interior, de volta ao líquido extracelular. Logo, a bomba de sódio e potássio bombeia ativamente o sódio para fora e o potássio para dentro na proporção de três íons sódio para dois íons potássio.



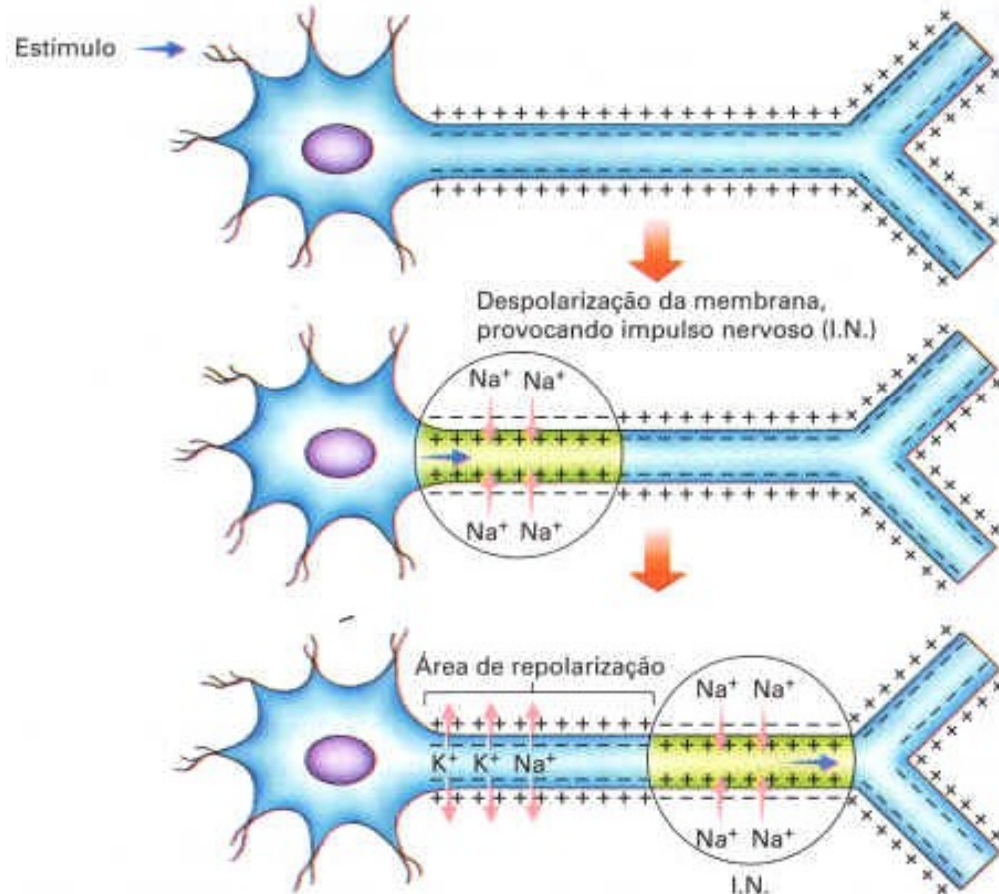
Comportamento Elétrico: Concentrações Iônicas

- Como a saída de sódio é maior que a entrada de potássio, cria-se uma diferença de cargas elétricas entre os meios intra e extracelular: há déficit de cargas positivas dentro da célula e as faces da membrana mantêm-se eletricamente carregadas, com potencial menor no interior da célula.
- O potencial eletronegativo criado no interior da fibra nervosa devido à bomba de sódio e potássio é chamado potencial de repouso da membrana, ficando o exterior da membrana positivo e o interior negativo, deixando a membrana polarizada.

Comportamento Elétrico: Concentrações Iônicas

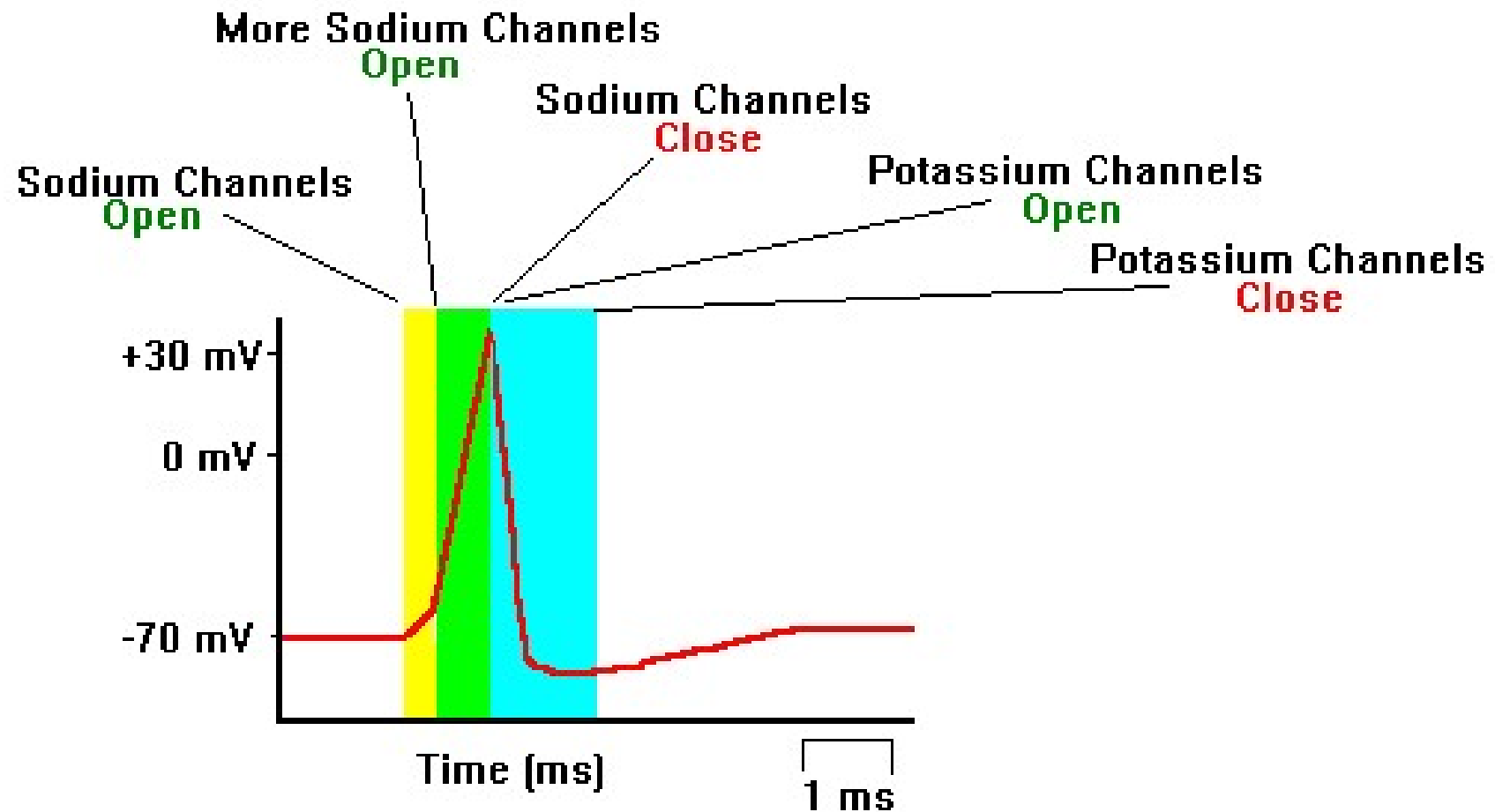
- Ao ser estimulada, uma pequena região da membrana torna-se permeável ao sódio (abertura dos canais de sódio). Como a concentração desse íon é maior fora do que dentro da célula, o sódio atravessa a membrana no sentido do interior da célula. A entrada de sódio é acompanhada pela pequena saída de potássio. Esta inversão vai sendo transmitida ao longo do axônio, e todo esse processo é denominado onda de despolarização.

Comportamento Elétrico: Concentrações Iônicas



Os impulsos nervosos ou potenciais de ação se devem à despolarização da membrana além de seu nível crítico que quando ultrapassado causa o disparo do potencial de ação. Este é semelhante em tamanho e duração para cada neurônio e não diminui à medida que é conduzido ao longo do axônio (tem tamanho e duração fixos). A despolarização crescente de um neurônio não tem qualquer efeito até que se cruze o limiar, quando surge o potencial de ação.

Comportamento Elétrico: Concentrações Iônicas

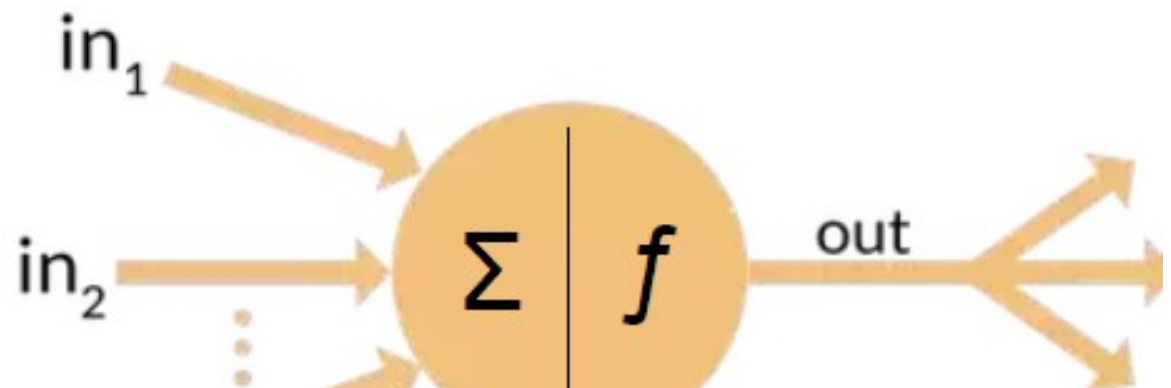
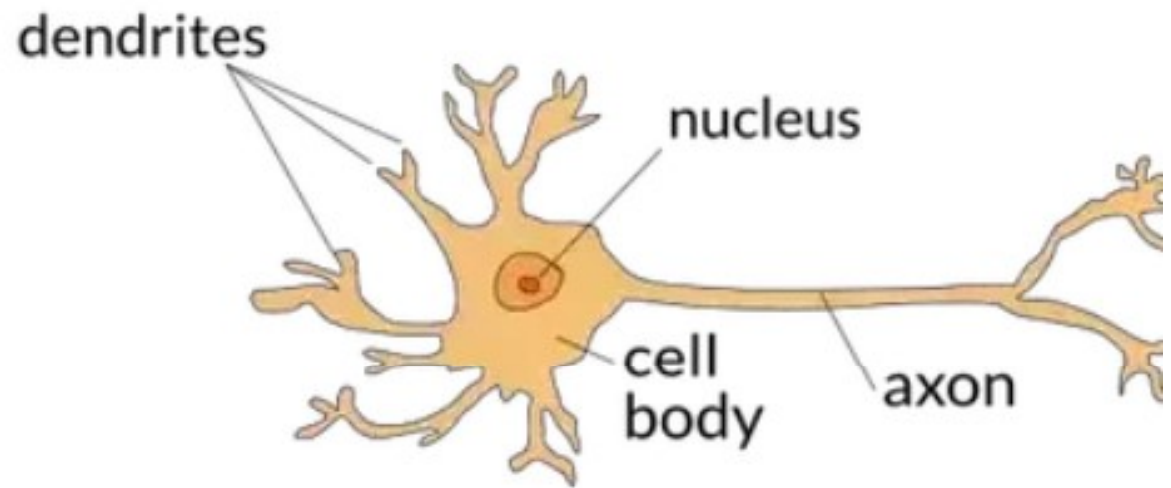


Comportamento Elétrico: Concentrações Iônicas

- Quando o neurônio é suficientemente estimulado ele produz um potencial de ação.
- Se o estímulo ultrapassa um limiar, os canais de sódio abrem-se repentinamente, levando sódio para dentro da célula. Esta dinâmica leva a despolarização local.
- O neurônio retorna ao repouso após o fechamento dos canais sódio e os canais de potássio abrem-se brevemente.
- A amplitude de um potencial de ação é sempre a mesma, não dependendo da intensidade do estímulo, dependendo apenas de ultrapassar o valor do limiar.
- Depois da despolarização, ocorrer um curto período refratário caracterizado por não permitir a ocorrência de um outro potencial de ação.

Arcabouço Comum às Redes Neurais

- Aracabouço Geral:
O objetivo é montar um modelo de neurônio e rede deles inspirado em neurônios reais. Esta abstração apresenta alguns fatores comuns que serão tratados a seguir:



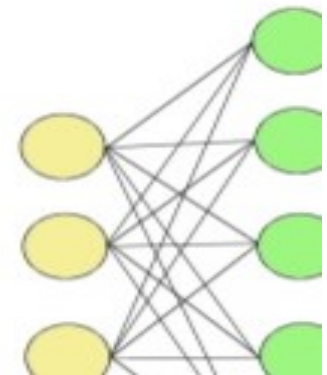
Arcabouço Comum às Redes Neurais

- Cada elemento da rede neural é chamado de nodo;
- As unidades são conectadas por ligações ou conexões (*links*);
- Cada conexão tem um valor numérico, o peso sináptico;
- Estes componentes definem a arquitetura de uma rede neural.

node called **unit**.

links.

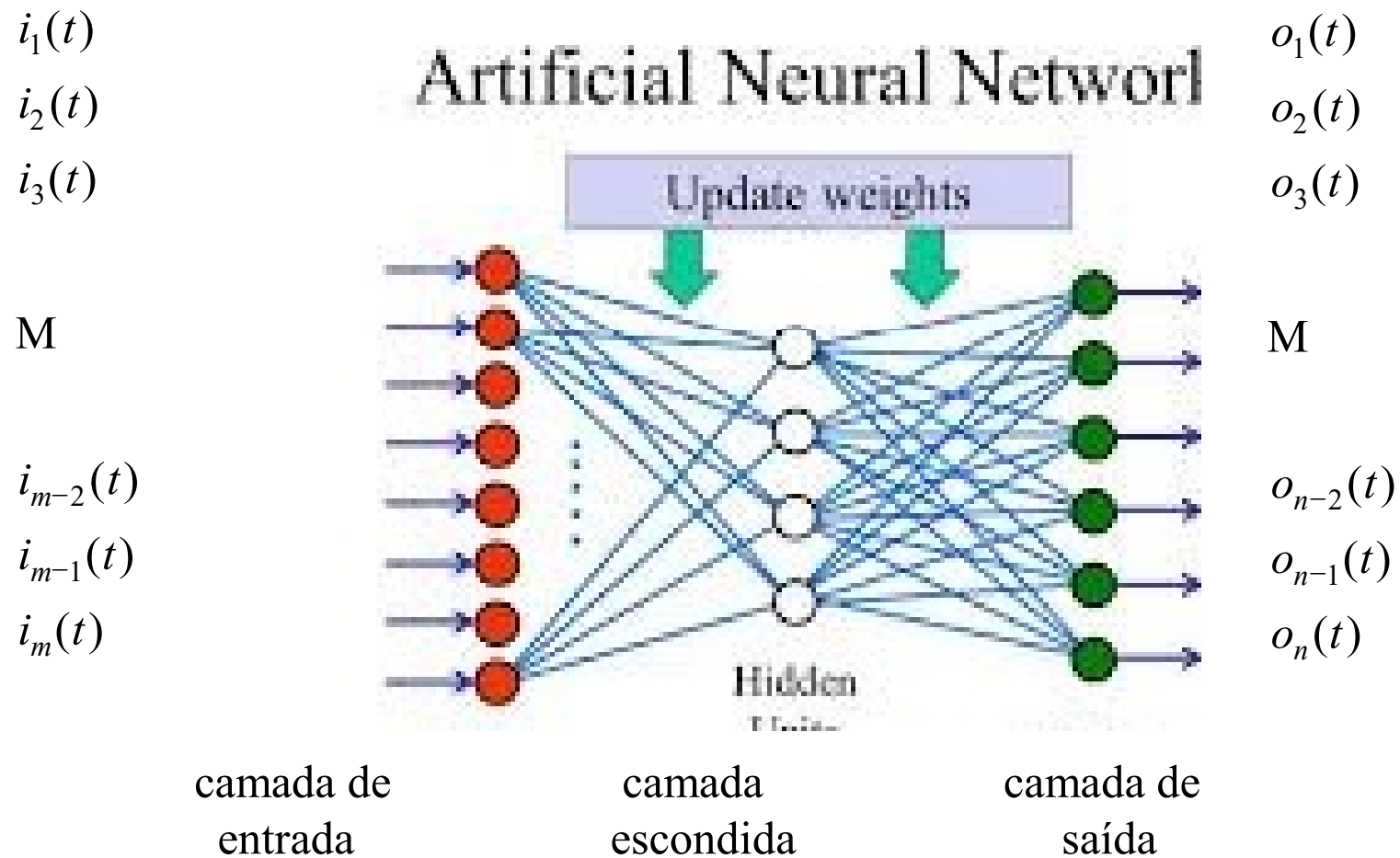
weight



Arcabouço Comum às Redes Neurais

- Arcabouço (Framework) Geral: Esta é uma proposta de um número de componentes comuns a muitos dos modelos de redes neurais. Rumelhart e colegas (1988) apontam oito componentes fundamentais:
 - Um conjunto de unidades de processamento;
 - Definição de diferentes estados de ativação;
 - Uma função de saída para cada unidade;
 - Um padrão de conectividade entre as diferentes unidades;
 - Uma regra de propagação;
 - Uma regra de ativação;
 - Uma regra de aprendizagem;
 - Uma representação do meio ambiente.

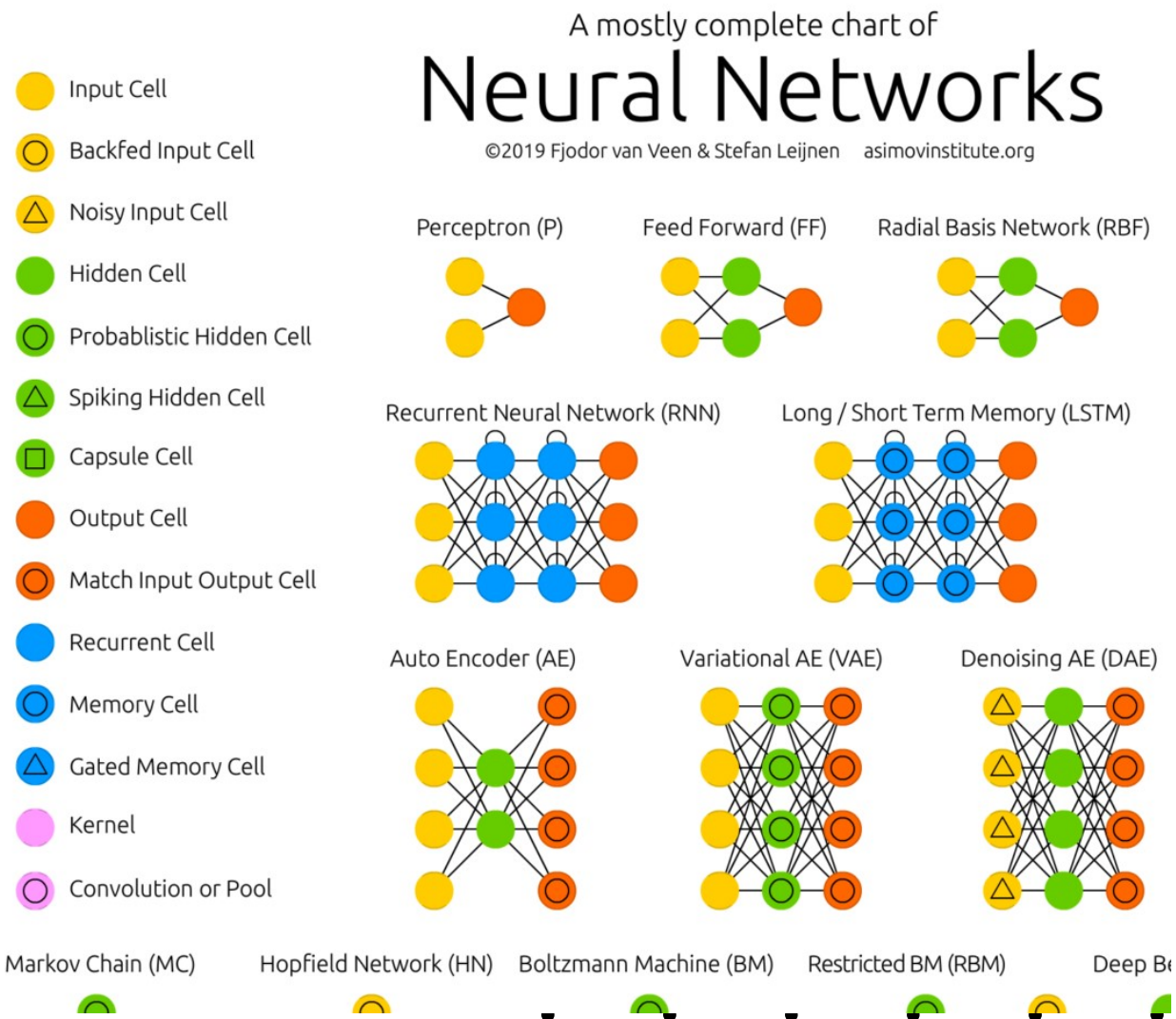
Topologia de uma Rede Neural Típica



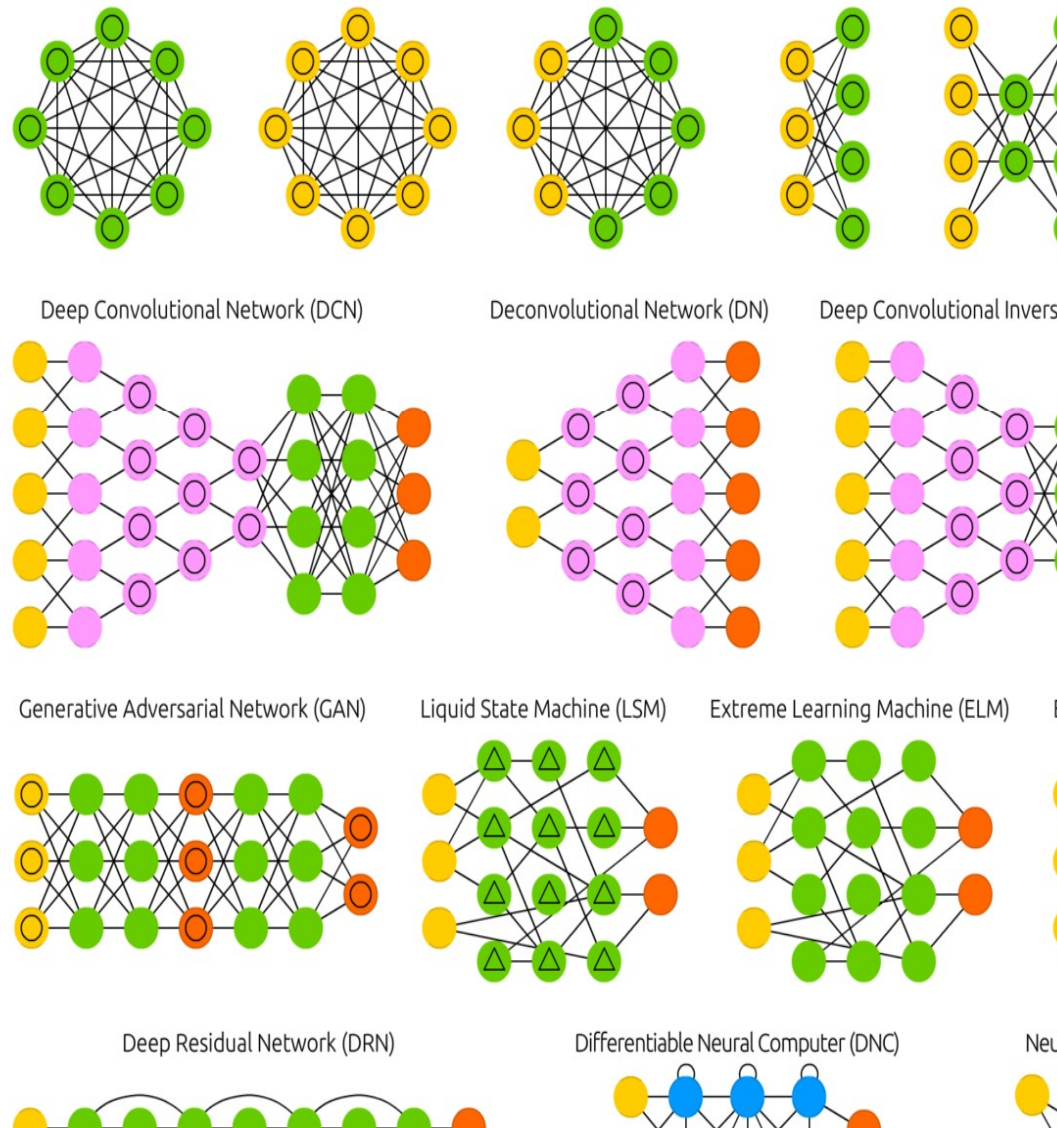
Topologia de uma Rede Neural Típica

- A figura mostra uma rede caracterizada por:
 - Três camadas: entrada, escondida e saída.
 - Existência de conexões entre todos os elementos de camadas vizinhas, não entre elementos da mesma camada ou de camadas não vizinhas.
 - Vetor de entrada $(i_1(t) \ i_2(t) \ ... \ i_{m-1}(t) \ i_m(t))^T$ apresentado à rede gera um vetor de ativação $(a_1(t) \ a_2(t) \ ... \ a_{m-1}(t) \ a_m(t))^T$ na camada de entrada. As saídas das unidades desta camada são transmitidas para a camada escondida que repete o processo anterior. A camada escondida, propaga sua saída com direção à camada de saída que produz um vetor de saída, que é dado por $(o_1(t) \ o_2(t) \ ... \ o_{n-1}(t) \ o_n(t))^T$.

Topologia de uma Rede Neural Típica



Topologia de uma Rede Neural Típica



Um Conjunto de Unidades de Processamento

- De início, determina-se o grau de semelhança entre uma unidade de processamento e um neurônio real. Em geral, são utilizadas abstrações dos neurônios, de sua arquitetura quando ligados em rede e das funções desempenhadas por esta.
- O segundo fator de interesse é o significado de cada unidade da rede: o que representam individualmente e em conjunto. Este é o problema da representação cujas formas mais comuns são:
 - Local: Cada unidade pode representar objetos conceituais particulares tais como características, letras, palavras, conceitos.
 - Distribuída: As unidades têm significado abstrato, logo o conjunto delas é que determina o significado de cada padrão.

Um conjunto de Unidades de Processamento

- Em resumo tem-se que:
 - *Representação distribuída*: Cada unidade representa entidades similares a pequenas características (microcaracterísticas), em geral sem significado próprio.
 - *Representação local*: Cada unidade tem significado próprio e representa um único conceito ou tem significado ainda mais geral.

Definição de Diferentes Estados de Ativação

- Cada estado de ativação de uma unidade de processamento tem um significado associado a ele. Tal estado pode ser especificado por um vetor de ativação no qual cada elemento representa o estado de ativação de uma unidade no instante de tempo t . É o padrão de ativação de um conjunto de unidades que captura o objeto ou evento que a rede representa em qualquer tempo.
- Existem algumas possibilidades de ativação das unidades. As escolhas levam em conta aquilo que o modelo representa. De maneira geral, o estado de ativação pode ser:
 - Analógico (reais) ou digital (binários, bipolares ou série de valores);
 - Contínua ou discreta;
 - Limitado ou ilimitado.

Definição de Diferentes Estados de Ativação

- Exemplos:
 - $[0,1]$ - ativação binária (digital).
 - $[0...1]$ - ativação analógica, limitada expressa em valores reais.
 - $\{1,2,...,9\}$ - ativação digital, limitada expressa em série de valores.
- Os diferentes estados de ativação objetivam representar o nível de atividade de um neurônio.

Uma Função de Saída

- Os sinais que fazem com que as unidades interajam são transmitidos por uma função de saída. A intensidade com que uma unidade interage sobre outra depende do nível de ativação da unidade em questão. A função de saída transforma a ativação da unidade de processamento em um sinal de saída: $o_i(t)=f_i(a_i(t))$, para a unidade i . Por exemplo,

$$o_i(t)=a_i(t);$$

$$o_i(t)=a_i(t), \quad \text{para } a_i(t)>T,$$

$$0, \quad \text{para } a_i(t)<T,$$

$$o_i(t-1), \quad \text{para } a_i(t)=T.$$

- A função de saída determina se uma unidade transmite informações para outras unidades e a intensidade desta transmissão.

Um Padrão de Conectividade

- O padrão de conectividade detém o conhecimento da rede neural. Ele também determina como tal rede responde a uma dada entrada. Este padrão é expresso por um conjunto de pesos representando conexões excitatórias ou inibitórias entre as unidades de processamento.
- Em geral, os diversos valores das conexões são reunidos em uma matriz de pesos ou matriz de conectividade (W) na qual cada componente representa a força da conexão entre as unidades u_i e u_j .
- Os valores das conexões, em geral, representam: conexão excitatória ($w_{ji} > 0$), conexão inibitória ($w_{ji} < 0$), e conexão desativada ($w_{ji} = 0$). O módulo de w_{ji} determina a intensidade da conexão.
- O padrão de conectividade representa as sinapses. O conhecimento está codificado nestas conexões.

Uma Regra de Propagação

- A função desta regra é combinar os sinais de entrada de uma unidade de processamento com os valores da matriz de conexões para produzir o efeito total “sentido” por esta unidade. O valor propagado para a unidade j é denotado como net_j .
- Por exemplo, $net_j = \sum_{i=1}^n w_{ji} o_i$
onde o_i é uma entrada da unidade j vinda da unidade i e w_{ji} é o valor da conexão entre u_i e u_j .
- A regra de propagação equivale a modificação da situação eletroquímica de um neurônio.

Uma Regra de Ativação

- Esta regra produz um novo estado de ativação da unidade de processamento j no qual se considera o net_j desta unidade e seu estado de ativação presente.
- Para uma rede com estados de ativação discretos, a regra de ativação pode ser definida como

$$a_j(k+1) = g_j(a_j(k), net_j(k))$$

- Para uma rede com estados de ativação contínuos, a regra de ativação pode ser definida como

$$a_j(t+\Delta t) = h_j(a_j(t), net_j(t))$$

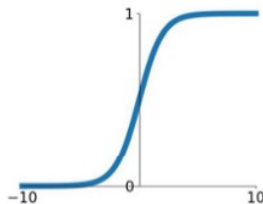
Uma Regra de Ativação

- Alguns exemplos de regras de ativação:

Activation Functions

Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



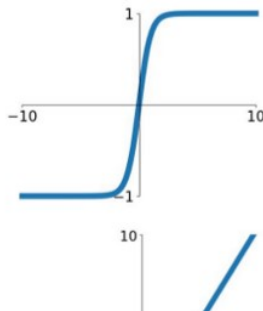
Leaky ReLU

$$\max(0.1x, x)$$

-16

tanh

$$\tanh(x)$$



Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

ReLU

ELU

- A regra de ativação representa o nível de atividade de uma unidade de processamento.



Uma Regra de Aprendizagem

- A mudança dos valores das conexões envolve o processo de aprendizagem. As conexões podem ser modificadas de três modos: aparecimento de novas conexões, perda ou modificação de conexões existentes. Os dois primeiros modos são casos particulares do terceiro pois em qualquer representação, conexões podem se tornar ativas ou inativas por treinamento.
- A magnitude da conexão é modificada através da experiência. Em geral, regras de aprendizagem são baseadas na regra de Hebb (1949). Nesta regra, a idéia básica é: ‘se duas unidades têm conexões entre si e estas unidades estão ativas, o valor da conexão entre elas cresce.’

Uma Regra de Aprendizagem

- Exemplos:
 - Regra de Hebb: $w_{ji} = x_i * x_j$
 - Regra do perceptron: $\sum \sum (E_{ij} - O_{ij})$
 - Regra delta: $\Delta W = (d - y) x_i$
 - Regra de correlação: $\Delta W_{ij} = \eta x_i$
- As mudanças dos valores de conexões representam as modificações das condições nas sinapses;

Uma Representação do Meio Ambiente

- Cada problema exige a representação do meio ambiente onde o processo a ser modelado ocorre. Tal representação precisa ser correta e apropriadamente modelada. Deve ficar claro, a constituição do sistema representado pela rede e do meio ambiente. Além disto, deve ser criteriosamente definido como acontece a interação entre ambos.

Operação de uma Rede Neural

- Em geral estes modelos têm duas fases básicas: fase de treinamento ou aprendizagem e fase de funcionamento ou teste.
 - No treinamento a rede incorpora o conhecimento extraído das regularidades de uma massa de dados.
 - Na fase de funcionamento, a rede responde a estímulos do meio ambiente, dela própria ou de outras redes, utilizando o conhecimento previamente acrescentado.
- Uma RN aprende as transformações que levam objetos em seu espaço de entradas para objetos em seu espaço de saídas.

Operação de uma Rede Neural

- O treinamento e funcionamento de uma RN compreende as seguintes etapas:
 - (1) Apresentação das entradas;
 - (2) Propagação dos estímulos ao longo da rede;
 - (3) Atualização das saídas da rede;
 - (4) Adaptação dos valores das conexões entre unidades de processamento.
- A última etapa está presente apenas no treinamento, as demais estão presentes tanto na fase de treinamento quanto na de funcionamento.

Metodologia para Utilização de Redes Neurais

- Esta metodologia apresenta sete passos:
 - (1) Levantamento do conjunto de dados relativo ao problema;
 - (2) Pré-processamento dos padrões selecionados;
 - (3) Organização dos padrões escolhidos em conjuntos;
 - (4) Definição da rede neural;
 - (5) Treinamento da rede neural;
 - (6) Teste da rede neural;
 - (7) Integração da rede neural ao meio ambiente.
- Em seguida discute-se cada um destes itens separadamente.

Levantamento do Conjunto de Dados Relativo ao Problema

- O primeiro passo para propor e implementar um sistema utilizando redes neurais consiste no levantamento das informações para treinamento e teste do sistema. Assim, deve-se usar a quantidade de informações adequada, coerentes com a disponibilidade, recursos e o tipo do problema. É importante salientar que qualquer quantidade de dados deve caracterizar tipicamente o domínio em estudo.
- A qualidade dos dados coletados é mais importante que sua quantidade (Bailey e Thompson, 1990). Portanto, os dados selecionados devem ser significativos, cobrir todo o domínio do problema, ter erros e ambigüidades minimizadas. Ferramentas estatísticas são empregadas na seleção das variáveis mais significativas do processo em estudo.

Pré-processamento dos Padrões Seleccionados - I

- Antes do pré-processamento propriamente dito, analisa-se os padrões consideradas no processo para escolha da estratégia de pré-processamento. Stein (1993) recomenda a utilização de histogramas que permitem visualização da distribuição de cada padrão envolvido. O pré-processamento consiste da preparação da representação dos padrões envolvidos no processo para que a rede os “entenda” e processe. Os padrões dão origem a variáveis que alimentarão a rede neural. Estas variáveis podem ser classificadas em dois tipos: nominais e numéricas. Elas são escolhidas em sintonia com os padrões que se quer processar.

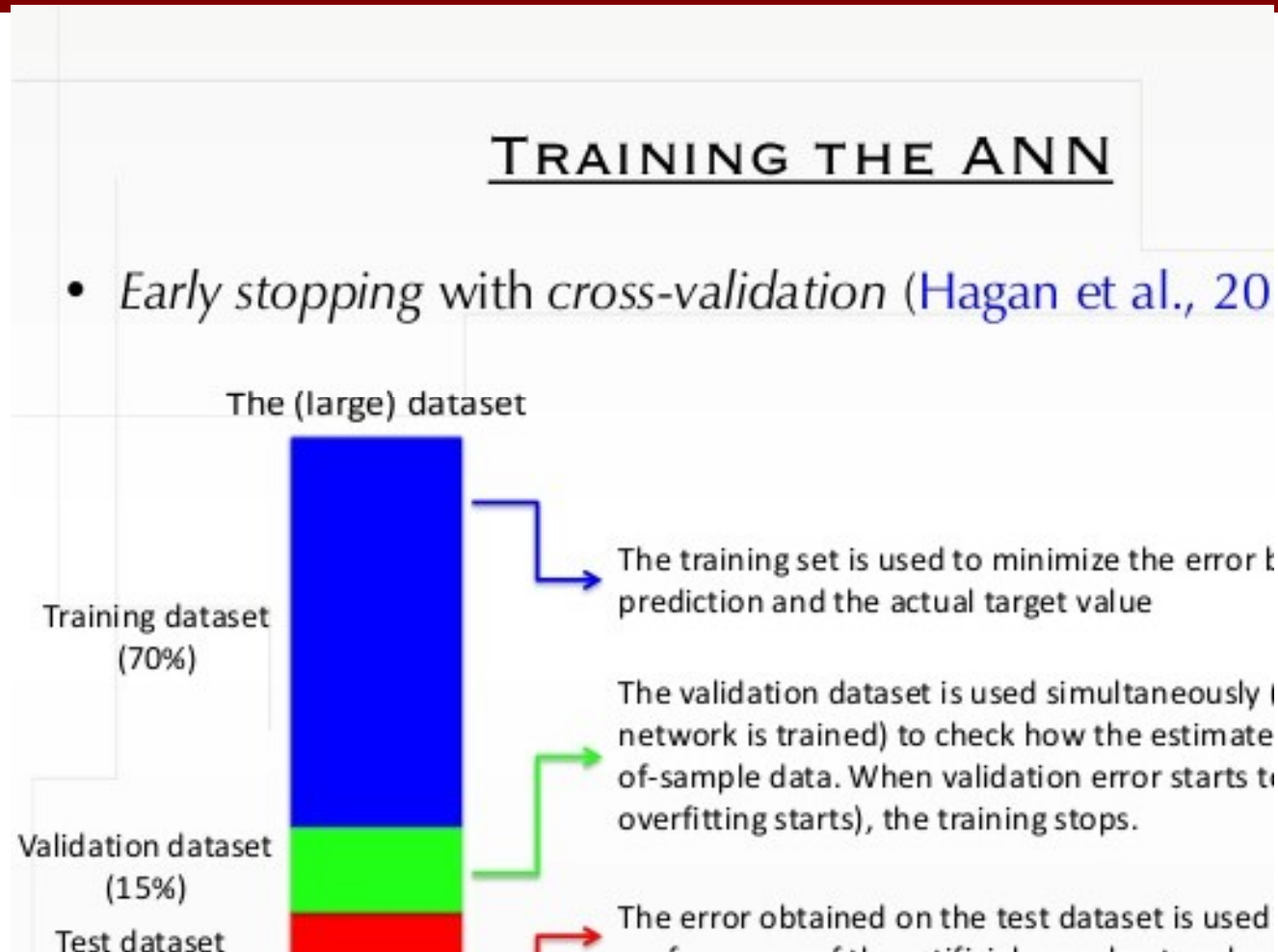
Pré-processamento dos Padrões Seleccionados - II

- As variáveis nominais são representadas por vetores de dois estados (binários ou bipolares) e se referem a classes com valores de ativação mutuamente excludentes. As variáveis nominais podem também representar variáveis numéricas que representam códigos numéricos tais como código de endereçamento postal ou identificadores numéricos de objetos. Em geral, as variáveis numéricas devem ser escalonadas em intervalos convenientes para o problema. Usualmente se escala estas variáveis entre zero e um.
- O pré-processamento também trata de outras questões: filtragem de dados para redução de ruídos, avaliação de valores médios para utilização posterior, complementação de informações ausentes nos padrões de treinamento e aproximação ou arredondamento de valores.

Organização do Padrões Escolhidos em Conjuntos

- Os dados selecionados e tratados são, em geral, separados em três conjuntos: treinamento, validação e teste. Os dados de treinamento são empregados para ajustar os pesos da rede durante sua fase de aprendizagem. Ainda nesta etapa, os dados de validação são utilizados para verificar a capacidade de generalização da rede. Os dados de teste são usados para avaliar o desempenho da rede resultante em condições de operação. Isto acontece durante a fase de teste. Todos estes conjuntos são usualmente apresentados de maneira aleatória à rede.

Organização do Padrões Escolhidos em Conjuntos



Definição da Rede Neural

- Este passo visa definir a configuração da rede, englobando: (1) Seleção do paradigma neural; (2) determinação da topologia da rede; (3) estabelecimento das funções da rede.
- Estas definições são normalmente encontradas por tentativa e erro. Também dependem da experiência do projetista, embora alguns autores apontem caminhos que auxiliam nesta tarefa (Caudil, 1991). O paradigma neural é selecionado de acordo com o problema que se está atacando. Logo, em geral, a escolha recai em um paradigma que tenha sido testado em problemas similares ou numa mesma família de problemas daquele em questão. A determinação da topologia inclui a especificação do número de unidades, seu arranjo espacial e o tipo de conexão entre unidades. As funções da rede compreendem as regras de propagação, ativação, saída e treinamento e seus parâmetros.

Treinamento da Rede Neural - I

- Esta etapa consiste do ajuste dos parâmetros livres (pesos, limiares, centros de bacias, etc) que compõem o mapeamento entrada-saída da rede. Nesta etapa o projetista deve prestar atenção na inicialização da rede, na sua estratégia de treinamento e no tempo de treinamento.
- A rede é normalmente inicializada com valores limitados e aleatórios ou valores pré-fixados que denotam o conhecimento prévio da rede. O aprendizado pode ser de dois modos: modo “on-line” ou padrão e modo “batch”. No modo padrão a correção de pesos acontece a cada apresentação de um novo padrão à rede. No modo batch, todos os padrões de treinamento são apresentados à rede e sua adaptação acontece como função do erro médio dos padrões. A escolha do método depende do problema em estudo.

Treinamento da Rede Neural - II

- O tempo de aprendizagem depende de fatores tais como algoritmo escolhido, complexidade do problema, entre outros. Este tempo também depende do critério de parada do modelo. Em geral os critérios de parada são função do erro das respostas da rede, da variação destas respostas para mesmas entradas, da capacidade de generalização da rede ou do número máximo de ciclos permitidos à rede.

Teste e Integração da Rede Neural

- Teste: Esta etapa visa medir o desempenho da rede, utilizando padrões não empregados na fase de treinamento.
- Integração: A integração da rede treinada e testada ao meio ambiente demanda interface adequada, capacidade de aquisição de dados do meio e capacidade de atuar no meio através de suas saídas.

Bibliografia

- Anderson, J. A. (1995). *An Introduction to Neural Networks*. Cambridge: The MIT Press.
- Beale, R. and Jackson, T. (1990). *Neural Computing: An Introduction*. Adam Hilger.
- Hassoun, M. H. (1995). *Fundamentals of Artificial Neural Networks*. Cambridge: The MIT Press.
- Haykin, S. (2008). *Neural Networks and Learning Machines*. Third Edition. Pearson.
- Rumelhart, D. E. and McClelland, J. L. (1988). *Parallel Distributed Processing*, 1, 2, and 3. Cambridge, MA: MIT Press.