TIVIDADE DE IICI4

Fazer uma pesquisa sobre rede neural com base nas seguintes questões a seguir:

* O que são redes neurais?

Resp.:

Redes neurais são sistemas de computação com nós interconectados que funcionam como os neurônios do cérebro humano. Usando algoritmos, elas podem reconhecer padrões escondidos e correlações em dados brutos, agrupá-los e classificá-los, e – com o tempo – aprender e melhorar continuamente

* Onde as redes podem ser aplicadas no dia a dia?

Resp.:

As redes neurais também são idealmente desenvolvidas para ajudar as pessoas a resolver problemas complexos em diversas situações da vida real. Elas podem aprender e modelar relações entre entradas e saídas de dados que são não-lineares e complexos; realizar generalizações e inferências; revelar relacionamentos, padrões e predições ocultas e modelar dados altamente voláteis (como dados de séries temporais financeiras) e variâncias necessárias para prever eventos raros (como detecção de fraudes). Como resultado, as redes neurais podem melhorar processos de decisão em diversas áreas, como:

* Detecção de fraude em cartões de crédito e assistência médica;
* Otimização de logística para redes de transporte;
* Reconhecimento de caracteres e de voz, também conhecido como processamento de linguagem natural;
* Diagnósticos médicos;
* Marketing direcionado;
* Predições financeiras de ações de mercado, moeda, opções, futuros, falência e classificação de títulos;
* Sistemas de controle robóticos;
* Previsão de carga elétrica e demanda de energia;
* Controle de qualidade e de processos;
* Identificação de compostos químicos;
* Avaliação do ecossistema;
* Visão computacional para interpretar fotos e vídeos não-tratados (por exemplo, na obtenção de imagens médicas, robótica e reconhecimento facial).

* O são treinamentos, ativação e generalização?

Resp.:

* Treinamento: Treinar uma rede neural como a de Hopfield por exemplo envolve abaixar a energia dos estados que a rede se “recorde”. Isto permite que a rede sirva como um sistema da memória endereçável satisfeita, aquele é dizer, a rede convergirá a um estado “recordado” se for dado somente a parte do estado. A rede pode ser usada recuperar de uma entrada distorcida o estado treinado que é o mais similar a essa entrada. Isto é chamado memória associativa porque recupera memórias na base da similaridade. Por exemplo, se nós treinarmos uma rede de Hopfield com cinco unidades de modo que o estado (1, 0, 1, 0, 1). seja um mínimo da energia, e nós damos à rede o estado (1, 0, 0, 0, 1) que convergirá (1, 0, 1, 0, 1). Assim, a rede está treinada corretamente quando a energia dos estados que a rede deve recordar é mínimos locais. O treinamento é de suma importância para o desempenho eficaz do mesmo.

* Ativação: Logo após o somatório ponderado dos neurônios, o valor obtido passa por mais um cálculo, chamado de função de ativação ou transferência. Seu objetivo é limitar a amplitude de saída do neurônio, ou seja, o valor obtido no somatório é normalizado dentro de um intervalo fechado, como [0,1], podendo ser interpretado também como a probabilidade do resultado.

Existem diversas funções lineares como a função degrau e identidade e não lineares como a gaussiana, tangente hiperbólica, sigmoide, entre outras, que testam o potencial de ativação através de sua fórmula. Em funções lineares, uma das técnicas mais populares é a função degrau, que determina, através de condições, o resultado final. Ou seja, é possível determinar que se o resultado da soma ponderada for menor ou igual zero, o algoritmo entregará um resultado, caso seja maior, ele entregará outro valor final. As condições dependem do que está se classificando.

Já em funções não lineares, a técnica logística sigmoidal é uma das mais populares em redes MLP e sua fórmula pode ser visualizada na imagem ao lado, no qual o símbolo *x*representa o valor obtido na soma ponderada do neurônio.

* Generalização: A generalização utiliza uma subdivisão do conjunto de treinamento, criando um conjunto de validação, utilizado para verificar a eficiência da rede quanto a sua capacidade de generalizar ou abstrair (interpretar com maior eficiência e tolerância aos dados de treinamento a partir da ativação), podendo ser empregado como critério de parada do treinamento.

* O que são redes supervisionadas e não supervisionadas?

Resp.:

Conceitualmente toda rede neural necessita de uma forma de aprendizado.

O aprendizado consiste na modificação dos pesos das conexões entre os neurônios, os pesos iniciais (sinapses) são modificados de forma iterativa, por um algoritmo que segue um dos seguintes paradigmas de aprendizado sendo eles:

* Aprendizado Supervisionado: é apresentado um conjunto de treino, consistindo de entradas e correspondentes saídas desejadas.
* Aprendizado Não-supervisionado: A rede atualiza seus pesos sem o uso de pares entrada-saídas desejadas e sem indicações sobre a adequação das saídas produzidas.

Existe ainda o paradigma do Aprendizado por Reforço onde para cada entrada apresentada, é produzida uma indicação (reforço) sobre a adequação das saídas correspondentes produzidas pela rede.

* Explique o que é neurônio artificial?

Resp.:

Pensando em como os neurônios trabalham, pesquisadores desenvolveram neurônios artificiais. Cada um tem dois ou mais receptores de entrada, responsáveis por perceberem um determinado tipo de sinal. Eles também possuem um corpo de processadores, responsável por um sistema de feedback que modifica sua própria programação dependendo dos dados de entrada e saída. Finalmente, eles possuem uma saída binária para apresentar a resposta “Sim” ou “Não”, dependendo do resultado do processamento.

Um neurônio artificial é capaz de um único processamento. Porém, ligar vários neurônios similares em rede, faz com que o sistema consiga processar mais informações e oferecer mais resultados.

Por exemplo, é possível criar um sistema para identificação de bananas e maçãs. Para tal, cria-se neurônios sensíveis a cor e forma. Os de cor percebem o amarelo e o vermelho. Os de forma percebem o redondo e o comprido. Cada neurônio, então, possui quatro entradas, uma para cada informação.

Para obter um melhor rendimento do sistema, cria-se uma rede em camadas: uma primeira camada com quatro neurônios (um para cada sinal de entrada), uma segunda camada oculta de processamento com três neurônios e uma camada de saída com dois neurônios, um para avisar quando é uma maça e outro para avisar quando é uma banana.

* O que é BIAS e para que serve?

Resp.:

Bias é uma entrada de valor "1" associada a um peso "b" em cada neurônio. Sua função é aumentar ou diminuir a entrada líquida, de forma a transladar a função de ativação no eixo.

Exemplo:

Para aproximar um conjunto de pontos a uma reta, usamos y = a\*x + b\*1, onde a e b são constantes. x eh uma entrada associada a um peso a e temos um peso b associado a entrada 1.

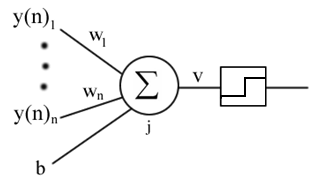
Agora imagine que a função de ativação da rede é uma função linear.

* Explique sobre a rede perceptron e a rede Adalaine, apresentando a diferença entre elas.

Resp.:

* **Perceptron**: Rede Neural Artificial *Perceptron* foi introduzida por Frank Rosenblatt, inspirado nos trabalhos de Walter Pitts e Warren Sturgis McCulloch. Esse modelo é um dos mais antigos e lida com um único neurônio, classificando o resultado de forma linear.

Na figura abaixo, o neurônio artificial é um Perceptron que recebe diversos valores de entradas *y(n)*. Essas entradas multiplicam-se pelo peso da sinapse *w* e, no final, somam-se formando um conjunto de entrada *ξ* = ∑ *w \* y(n)*. Esse resultado passa por uma função de ativação linear (que será explicada detalhadamente mais à frente desse artigo) e transmite a saída *v*. Quando o valor *ξ* exceder o limite da função de ativação, o neurônio será ativado e retornará um valor.

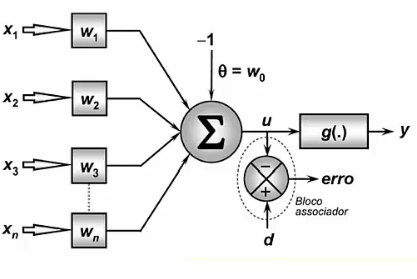


* **Adaline:** A rede Adaline (Adaptive Linear Element) proposta por Widrow e Hoff (1960) tem a mesma estrutura do Perceptron, diferenciando apenas no algoritmo de treinamento.

Características da Adaline:

* Rede adaptativa
* Inclusão de um algoritmo pioneiro para treinamento de redes de múltiplas camadas, o algoritmo de treinamento da Regra Delta
* O algoritmo da Regra Delta é baseado no método dos mínimos quadrados.  
   Método de aprendizado mais suave.

A figura abaixo ilustra uma rede Adaline com *n* sinais de entrada (*x1…xn*) e uma única saída (*y*). Na Adaline é proposto a inclusão de um bloco de sinal de erro com o intuito de modificar os valores dos pesos antes de se submeter a função de ativação (*g*). Note também que ilustra apenas um único neurônio.



De modo semelhante ao Perceptron o processamento se dá pelas seguinte expressões:

wj•Xi 
izi 

Onde ***xi*** são os sinais de entrada, ***wi*** são os pesos sinápticos associados, **θ** é o *bias*(limiar de ativação), ***g(.)*** é a função de ativação (degrau ou degrau bipolar) e ***u*** é o potencial de ativação. O bloco associador demonstrado no diagrama acima tem uma função auxiliar no processo de treinamento da rede.

**erro = d-u**

Erro é igual ao valor esperado menos o valor do potencial de ativação produzido pela rede.

* Explique o que é a taxa de aprendizagem?

Resp.:

A taxa de aprendizado é uma constante de proporcionalidade no intervalo [0,1], pois este procedimento de aprendizado requer apenas que a mudança no peso seja proporcional.

Entretanto, o verdadeiro gradiente descendente requer que sejam tomados passos infinitesimais. Assim quanto maior for essa constante, maior será a mudança nos pesos, aumentando a velocidade do aprendizado, o que pode levar à uma oscilação do modelo na superfície de erro. O ideal seria utilizar a maior taxa de aprendizado possível que não levasse à uma oscilação, resultando em um aprendizado mais rápido.

<http://conteudo.icmc.usp.br/pessoas/andre/research/neural/MLP.htm>

* Procure e explique o algoritmo da rede perceptron.

Resp.:

Para a explicação do algoritmo da rede perceptron, em primeiro momento será ignorada a adição de um bias ao resultado do somatório dos inputs sem prejuízos do entendimento.

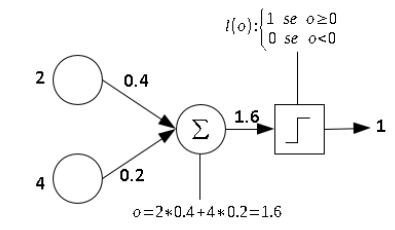
No algoritmo temos inputs representados por "X1" e "X2" e os respectivos pesos de cada input "W1" e "W2". Esses inputs passam por uma combinação linear ponderada por seus respectivos pesos, onde "O" é o output desse primeiro node. "O" então é avaliado por uma função de perda, para problemas de classificação binária é muito comum o uso da "0-1 loss function":

Ex.: 1 se O>=0

l(O)= [

0 se O <=0

Um exemplo de aplicação:



* Use o algoritmo para resolver o problema da porta AND.

Resp.: