

Deep Learning Book

Em Português, Online e Gratuito

Capítulo 2 – Uma Breve História das Redes Neurais Artificiais



Para compreender onde estamos hoje, precisamos olhar para o passado e analisar como chegamos até aqui. Vejamos então Uma Breve História das Redes Neurais Artificiais.

O cérebro humano é uma máquina altamente poderosa e complexa capaz de processar uma grande quantidade de informações em tempo mínimo. As unidades principais do cérebro são os neurônios e é por meio deles que as informações são transmitidas e processadas. As tarefas realizadas pelo cérebro intrigam os pesquisadores, como por exemplo, a capacidade do cérebro de reconhecer um rosto familiar dentre uma multidão em apenas milésimos de segundo. As respostas sobre alguns enigmas do funcionamento do cérebro ainda não foram respondidas e se perpetuam até os dias de hoje. O que é conhecido sobre o funcionamento do cérebro é que o mesmo desenvolve suas regras através da experiência adquirida em situações vividas anteriormente.



Fig1 – Cérebro humano, a máquina mais fantástica que existe no Planeta Terra.

O desenvolvimento do cérebro humano ocorre principalmente nos dois primeiros anos de vida, mas se arrasta por toda a vida. Inspirando-se neste modelo, diversos pesquisadores tentaram simular o funcionamento do cérebro, principalmente o processo de aprendizagem por experiência, a fim de criar sistemas inteligentes capazes de realizar tarefas como classificação, reconhecimento de padrões, processamento de imagens, entre outras atividades. Como resultado destas pesquisas surgiu o modelo do neurônio artificial e posteriormente um sistema com vários neurônios interconectados, a chamada Rede Neural.

Em 1943, o neurofisiologista Warren McCulloch e o matemático Walter Pitts escreveram um artigo sobre como os neurônios poderiam funcionar e para isso, eles modelaram uma rede neural simples usando circuitos elétricos.

Warren McCulloch e Walter Pitts criaram um modelo computacional para redes neurais baseadas em matemática e algoritmos denominados lógica de limiar (threshold logic). Este modelo abriu o caminho para a pesquisa da rede neural dividida em duas abordagens: uma abordagem focada em processos biológicos no cérebro, enquanto a outra focada na aplicação de redes neurais à inteligência artificial.

Em 1949, Donald Hebb escreveu *The Organization of Behavior*, uma obra que apontou o fato de que os caminhos neurais são fortalecidos cada vez que são usados, um conceito fundamentalmente essencial para a maneira como os humanos aprendem. Se dois nervos dispararem ao mesmo tempo, argumentou, a conexão entre eles é melhorada.

À medida que os computadores se tornaram mais avançados na década de 1950, finalmente foi possível simular uma hipotética rede neural. O primeiro passo para isso foi feito por Nathaniel Rochester dos laboratórios de pesquisa da IBM. Infelizmente para ele, a primeira tentativa de fazê-lo falhou.

No entanto, ao longo deste tempo, os defensores das “máquinas pensantes” continuaram a argumentar suas pesquisas. Em 1956, o Projeto de Pesquisa de Verão de Dartmouth sobre Inteligência Artificial proporcionou um impulso tanto à Inteligência Artificial como às Redes Neurais. Um dos resultados deste processo foi estimular a pesquisa em IA na parte de processamento neural.

Nos anos seguintes ao Projeto Dartmouth, John von Neumann sugeriu imitar funções simples de neurônios usando relés telegráficos ou tubos de vácuo. Além disso, Frank Rosenblatt, um neurobiologista, começou a trabalhar no Perceptron. Ele estava intrigado com o funcionamento do olho de uma mosca. Grande parte do processamento feito por uma mosca ao decidir fugir, é feito em seus olhos. O Perceptron, que resultou dessa pesquisa, foi construído em hardware e é a mais antiga rede neural ainda em uso hoje. Um Perceptron de camada única foi útil para classificar um conjunto de entradas de valor contínuo em uma de duas classes. O Perceptron calcula uma soma ponderada das entradas, subtrai um limite e passa um dos dois valores possíveis como resultado. Infelizmente, o Perceptron é limitado e foi comprovado como tal durante os “anos desiludidos” por Marvin Minsky e o livro de Seymour Papert de 1969, *Perceptrons*.

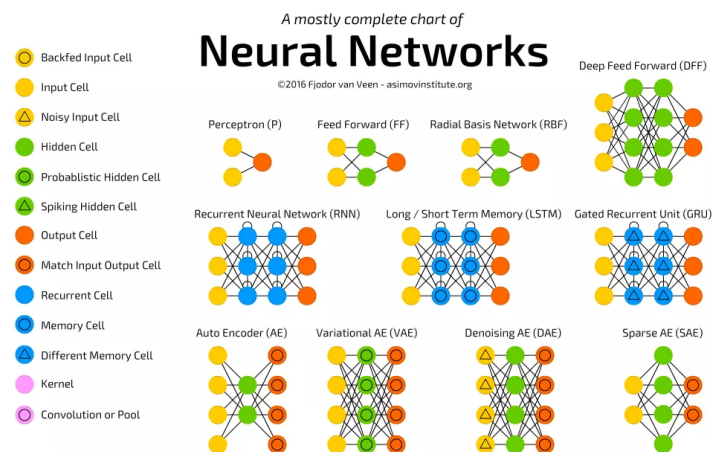


Fig2 – Algumas Arquiteturas de Redes Neurais

Em 1959, Bernard Widrow e Marcian Hoff, de Stanford, desenvolveram modelos denominados “ADALINE” e “MADALINE”. Em uma exibição típica do amor de Stanford por siglas, os nomes provêm do uso de múltiplos elementos ADaptive LINear. ADALINE foi desenvolvido para reconhecer padrões binários de modo que, se ele estivesse lendo bits de transmissão de uma linha telefônica, poderia prever o próximo bit. MADALINE foi a primeira rede neural aplicada a um problema do mundo real, usando um filtro adaptativo que elimina ecos nas linhas telefônicas. Embora o sistema seja tão antigo como os sistemas de controle de tráfego aéreo, ele ainda está em uso comercial.

Infelizmente, esses sucessos anteriores levaram as pessoas a exagerar o potencial das redes neurais, particularmente à luz da limitação na eletrônica, então disponível na época. Este exagero excessivo, que decorreu do mundo acadêmico e técnico, infectou a literatura geral da época. Muitas promessas foram feitas, mas o resultado foi o desapontamento. Além disso, muitos escritores começaram a refletir sobre o efeito que teria “máquinas pensantes” no homem. A série de Asimov em robôs revelou os efeitos sobre a moral e os valores do homem quando máquinas fossem capazes de fazer todo o trabalho da humanidade. Outros escritores criaram computadores mais sinistros, como HAL do filme 2001.

Toda essa discussão sobre o efeito da Inteligência Artificial sobre a vida humana, aliada aos poucos progressos, fizeram vozes respeitadas criticar a pesquisa em redes neurais. O resultado foi a redução drástica de grande parte do financiamento em pesquisas. Esse período de crescimento atrofiado durou até 1981, sendo conhecido como o Inverno da IA (AI Winter).

Em 1982, vários eventos provocaram um renovado interesse. John Hopfield da Caltech apresentou um documento à Academia Nacional de Ciências. A abordagem de Hopfield não era simplesmente modelar cérebros, mas criar dispositivos úteis. Com clareza e análise matemática, ele mostrou como essas redes poderiam funcionar e o que poderiam fazer. No entanto, o maior recurso de Hopfield foi seu carisma. Ele era articulado e simpático e isso colaborou bastante para que ele fosse ouvido.

Em 1985, o Instituto Americano de Física começou o que se tornou uma reunião anual – Redes Neurais para Computação. Em 1987, a primeira Conferência Internacional sobre Redes Neurais do Institute of Electrical and Electronic Engineer's (IEEE) atraiu mais de 1.800 participantes.

Em 1986, com redes neurais de várias camadas nas notícias, o problema era como estender a regra Widrow-Hoff para várias camadas. Três grupos independentes de pesquisadores, dentre os quais David Rumelhart, ex-membro do departamento de psicologia de Stanford, apresentaram ideias semelhantes que agora são chamadas de redes Backpropagation porque distribuem erros de reconhecimento de padrões em toda a rede. As redes híbridas utilizavam apenas duas camadas, essas redes de Backpropagation utilizam muitas. O resultado é que as redes de Backpropagation “aprendem” de forma mais lenta, pois necessitam, possivelmente, de milhares de iterações para aprender, mas geram um resultado muito preciso.

Agora, as redes neurais são usadas em várias aplicações. A ideia fundamental por trás da natureza das redes neurais é que, se ela funcionar na natureza, deve ser capaz de funcionar em computadores. O futuro das redes neurais, no entanto, reside no desenvolvimento de hardware. As redes neurais rápidas e eficientes dependem do hardware especificado para seu eventual uso.

O diagrama abaixo mostra alguns marcos importantes na evolução e pesquisa das redes neurais artificiais. O fato, é que ainda estamos escrevendo esta história e muita evolução está ocorrendo neste momento, através do trabalho de milhares de pesquisadores e profissionais de Inteligência Artificial em todo mundo. E você, não quer ajudar a escrever esta história?

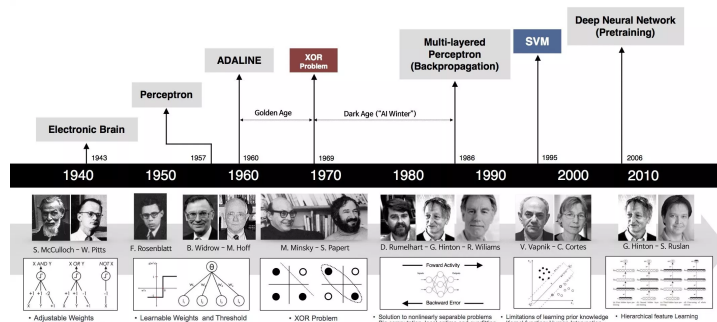


Fig3 – Marcos no desenvolvimento das redes neurais.

Podemos resumir assim os principais marcos na pesquisa e evolução das redes neurais artificiais até chegarmos ao Deep Learning:

1943: Warren McCulloch e Walter Pitts criam um modelo computacional para redes neurais baseadas em matemática e algoritmos denominados lógica de limiar.

1958: Frank Rosenblatt cria o Perceptron, um algoritmo para o reconhecimento de padrões baseado em uma rede neural computacional de duas camadas usando simples adição e subtração. Ele também propôs camadas adicionais com notações matemáticas, mas isso não seria realizado até 1975.

1980: Kunihiko Fukushima propõe a Neoconitron, uma rede neural de hierarquia, multicamada, que foi utilizada para o reconhecimento de caligrafia e outros problemas de reconhecimento de padrões.

1989: os cientistas conseguiram criar algoritmos que usavam redes neurais profundas, mas os tempos de treinamento para os sistemas foram medidos em dias, tornando-os impraticáveis para o uso no mundo real.

1992: Juyang Weng publica o Cresceptron, um método para realizar o reconhecimento de objetos 3-D automaticamente a partir de cenas desordenadas.

Meados dos anos 2000: o termo “aprendizagem profunda” começa a ganhar popularidade após um artigo de Geoffrey Hinton e Ruslan Salakhutdinov mostrar como uma rede neural de várias camadas poderia ser pré-treinada uma camada por vez.

2009: acontece o NIPS Workshop sobre Aprendizagem Profunda para Reconhecimento de Voz e descobre-se que com um conjunto de dados suficientemente grande, as redes neurais não precisam de pré-treinamento e as taxas de erro caem significativamente.

2012: algoritmos de reconhecimento de padrões artificiais alcançam desempenho em nível humano em determinadas tarefas. E o algoritmo de aprendizagem profunda do Google é capaz de identificar gatos.

2014: o Google compra a Startup de Inteligência Artificial chamada DeepMind, do Reino Unido, por £ 400m

2015: Facebook coloca a tecnologia de aprendizado profundo – chamada DeepFace – em operação para marcar e identificar automaticamente usuários do Facebook em fotografias. Algoritmos executam tarefas superiores de reconhecimento facial usando redes profundas que levam em conta 120 milhões de parâmetros.

2016: o algoritmo do Google DeepMind, AlphaGo, mapeia a arte do complexo jogo de tabuleiro Go e vence o campeão mundial de Go, Lee Sedol, em um torneio altamente divulgado em Seul.

2017: adoção em massa do Deep Learning em diversas aplicações corporativas e mobile, além do avanço em pesquisas. Todos os eventos de tecnologia ligados a Data Science, IA e Big Data, apontam Deep Learning como a principal tecnologia para criação de sistemas inteligentes.

A promessa do aprendizado profundo não é que os computadores comecem a pensar como seres humanos. Isso é como pedir uma maçã para se tornar uma laranja. Em vez disso, demonstra que, dado um conjunto de dados suficientemente grande, processadores rápidos e um algoritmo suficientemente sofisticado, os computadores podem começar a realizar tarefas que até então só podiam ser realizadas apenas por seres humanos, como reconhecer imagens e voz, criar obras de arte ou tomar decisões por si mesmo.

Os estudos sobre as redes neurais sofreram uma grande revolução a partir dos anos 80 e esta área de estudos tem se destacado, seja pelas promissoras características apresentadas pelos modelos de redes neurais propostos, seja pelas condições tecnológicas atuais de implementação que permitem desenvolver arrojadas implementações de arquiteturas neurais paralelas em hardwares dedicados, obtendo assim ótimas performances destes sistemas (bastante superiores aos sistemas convencionais). A evolução natural das redes neurais, são as redes neurais profundas (ou Deep Learning). Mas isso é o que vamos discutir no próximo capítulo! Até lá.

Referências:

Christopher D. Manning. (2015). Computational Linguistics and Deep Learning Computational Linguistics, 41(4), 701–707.

F. Rosenblatt. The perceptron, a perceiving and recognizing automaton Project Para. Cornell Aeronautical Laboratory, 1957.

W. S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics, 5(4):115–133, 1943.

The organization of behavior: A neuropsychological theory. D. O. Hebb. John Wiley And Sons, Inc., New York, 1949

B. Widrow et al. Adaptive “Adaline” neuron using chemical “memistors”. Number Technical Report 1553-2. Stanford Electron. Labs., Stanford, CA, October 1960.

“New Navy Device Learns By Doing”, New York Times, July 8, 1958.

Perceptrons. An Introduction to Computational Geometry. MARVIN MINSKY and SEYMOUR PAPERT. M.I.T. Press, Cambridge, Mass., 1969.

Minsky, M. (1952). A neural-analogue calculator based upon a probability model of reinforcement. Harvard University Psychological Laboratories internal report.

Compartilhe isso:

