

Nome: Lucas Henrique Mantovani Jacintho
nUSP: 10258942

O estudo das Redes Neurais se inicia aproximadamente em 1943, quando o neurofisiologista Warren McCulloch e o matemático Walter Pitts, devido aos avanços da eletrônica moderna, escreverem um artigo sobre o funcionamento dos neurônios humanos [1]. No artigo, modelaram o neurônio utilizando circuitos eletrônicos. O modelo é dividido em duas partes. A primeira parte recebe um estímulo (que pode ser entendido como uma entrada) e performa uma modificação do sinal (processamento), e a segunda parte é responsável pela tomada e transmissão da decisão [2].

Em 1949, Donald Hebb escreveu *The Organization of Behavior*, que apontava o aumento da força de conexão de neurônios à medida que eram utilizados, como um conceito fundamental para o aprendizado humano [4]. Esse conceito é um dos precursores das Redes Neurais, assim como a *lógica Threshold*. Esse último consiste no mapeamento de valores contínuos em valores discretos [5].

Já em 1950, houve a tentativa de traduzir essas redes para sistemas computacionais, feita por Nathaniel Rochester, membro dos laboratórios de pesquisa da IBM [5][4]. Porém, apenas em 1954 que a primeira rede foi implementada com sucesso, pelo MIT [5]. Na mesma época, a computação tradicional começou a florescer. Isso desacelerou o desenvolvimento das Redes Neurais, pois a pesquisa desse tema foi deixada em segundo plano [1].

Mesmo com a desaceleração, algumas pessoas ainda defendiam o desenvolvimento desse ramo da computação [1]. No ano de 1956, o *Dartmouth Summer Research Project on Artificial Intelligence* estimulou a pesquisa em ambos os ramos da inteligência: a inteligência artificial como conhecida pela indústria, e também em nível mais baixo, o processamento neuronal do cérebro [3].

No ano seguinte, John von Neumann sugeriu que as funções do neurônio fossem imitadas por relays de telégrafo ou tubos de vácuo [1]. Em 1958, Frank Rosenblatt iniciou os trabalhos que dariam origem ao Perceptron. O Perceptron foi modelado utilizando o neurônio de McCulloch-Pitts com a decisão sendo tomada após ser avaliada por um *Threshold linear*. Ou seja, após uma soma ponderada, o valor retornado seria '0' caso a soma estivesse abaixo do limiar escolhido, e '1' caso estivesse acima [5].

Posteriormente, em 1959, Bernard Widrow e Marcian Hoff desenvolveram os modelos "ADALINE" e "MADALINE". O primeiro modelo foi desenvolvido de forma a reconhecer padrões binários e, portanto, ao receber um bit, poderia prever o próximo. Já o segundo, foi o primeiro modelo a ser aplicado em um problema real, sendo utilizado com um filtro para eliminar o eco em chamadas telefônicas. Esse sistema permanece em uso até os dias atuais [4]. A diferença entre os modelos Perceptron e ADALINE está na forma do aprendizado. Enquanto o Perceptron utiliza as classes para aprender, o ADALINE utiliza os valores contínuos preditos para aprender os coeficientes do modelo. Ou seja, consegue obter uma "quantidade" de erro [6].

Após exagero do potencial das Redes Neurais na época, e promessas não cumpridas levaram ao declínio da popularidade do tema [1]. Junto com a publicação de Marvin Minsky, que argumentava ser impossível transformar o modelo Perceptron em um modelo multicamadas, as pesquisas em Redes Neurais ficaram praticamente paralisadas por pouco mais de uma década [5].

Foi apenas em 1982 que o interesse por Redes Neurais foi retomado na comunidade científica. O primeiro evento que contribuiu para essa retomada foi o paper apresentado por John Hopfield da Caltech para a *Academy of Sciences*. Sua abordagem era de criar dispositivos úteis, e não apenas modelar o funcionamento do cérebro humano. Com clareza e matemática simples, Hopfield demonstrou que as redes poderiam funcionar e do que elas seriam capazes [4].

Ainda em 1982, em uma conferência conjunta realizada por Japão e Estados Unidos, o Japão anunciou um esforço maior em Redes Neurais. Temendo ficar para trás, os Estados Unidos optaram por aumentar a destinação de fundos para pesquisas sobre o tema, contribuindo para a volta de sua popularidade na academia [1] [4].

Após sua segunda popularização, o tema passou a ser cada vez mais estudado e desenvolvido. Assim, diversas outras formas de modelos foram propostos. Podem ser citados, por exemplo, os modelos convolucionais, os modelos recorrentes, as Redes Neurais Adversariais, modelos generativos, entre diversos outros [7].

Os modelos convolucionais, como o nome sugere, são baseados na operação de convolução. Os modelos recorrentes, utilizam de conexões de uma camada com ela mesma, por exemplo, para manter a informação utilizada anteriormente, gerando um efeito memória [7]. As Redes Neurais Adversariais são modelos onde, no geral, há duas redes: uma rede que produz uma saída e outra que tenta classificar em conteúdo real ou artificial. Ambas as redes evoluem após "perderem" para a outra [8].

Devido ao aumento do poder computacional as capacidades das Redes Neurais também aumentaram. Porém, alguns dizem que o hardware já está próximo de seu limite físico, o que impactaria também nesse campo de estudo. De olho nesse possível problema, alguns pesquisadores já trabalham para desenvolver um hardware específico para ser utilizado por Redes Neurais, de modo a ter maior eficiência se comparado com um hardware de um computador de uso geral [9]. Destacam-se atualmente as GPUs e mais recentemente as TPUs [10].

Referências

1. <http://www2.psych.utoronto.ca/users/reingold/courses/ai/cache/neural4.html>
2. <https://towardsdatascience.com/mcculloch-pitts-model-5fdf65ac5dd1>
3. <https://medium.com/analytics-vidhya/brief-history-of-neural-networks-44c2bf72eec>
4. <https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/History/history1.html>
5. <https://towardsdatascience.com/a-concise-history-of-neural-networks-2070655d3fec>
6. <https://sebastianraschka.com/faq/docs/diff-perceptron-adaline-neuralnet.html>
7. <https://www.digitalvidya.com/blog/types-of-neural-networks/>
8. <https://towardsdatascience.com/understanding-generative-adversarial-networks-gans-cd6e4651a29>
9. https://en.wikipedia.org/wiki/AI_accelerator
10. https://en.wikipedia.org/wiki/Tensor_processing_unit