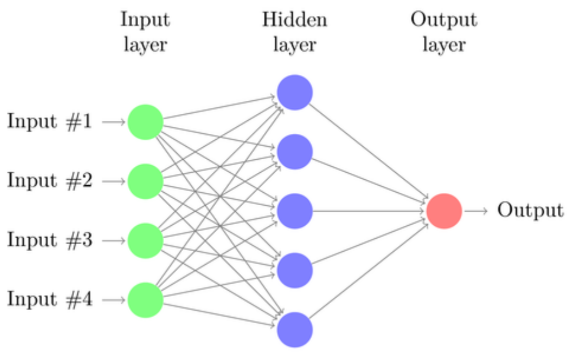
COCUS PORTUGAL
DATA SCIENTIST POSITION

Guilherme Giuliano Nicolau
PhD in Political Science by the University of São Paulo

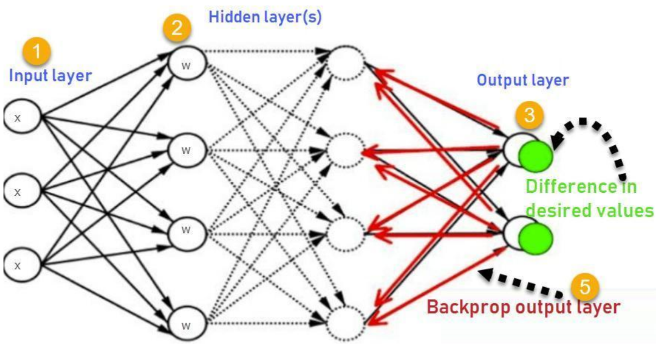
1. *Neural Network Explained*

Neural networks are machine learning architectures that reproduce the functioning of human neurons. They are not new and have been around since at least the 1960s, but they went through an 'ice age' until they reappeared in the 1990s, taking advantage of the greater computing power of our hardware. These are complex models capable of dealing with the non-linearity of our data and today already achieve predictive learning comparable to that of humans, not without high computational costs and carbon emissions.

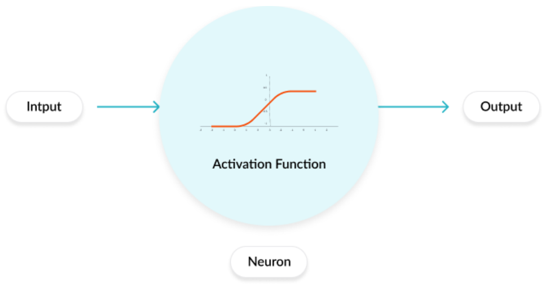
A simpler neural network consists of three layers of neurons, one for input, one hidden and one for output.



We always have an input layer and an output layer, but we can expand the hidden layers in order to increase the complexity of the model and improve its accuracy, constituting deep neural networks.

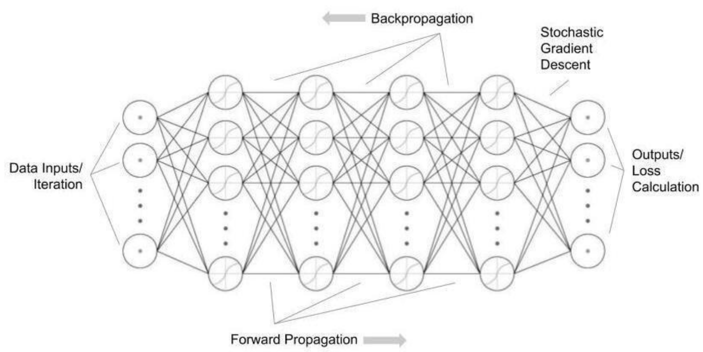


First we have a forward propagation. The input layer receives our data transformed into vectors and it will be multiplied by another vector of random weights within an activation function on a neuron.

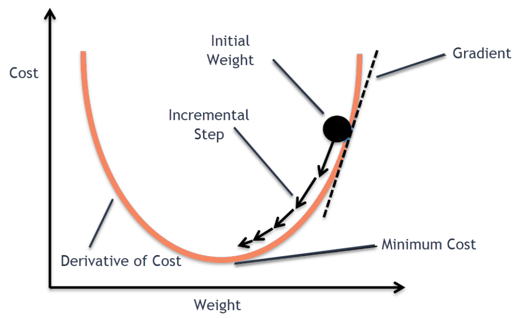


Our last layer always applies a specific activation function for our objective (e.g. softmax for classification problem), but each of our other deep layers can apply different activation functions. The last output layer returns all processed data.

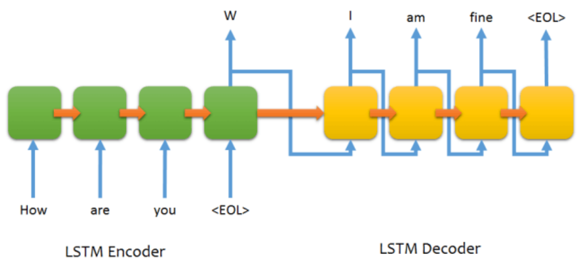
Once we have the output from the multiplicative combination between neurons and their hidden layers, we can calculate the error (loss) of the prediction and adjust the weights to compensate for it, thus repeating the process until a learning process is capable of a satisfactory predictive accuracy. This adjustment process can be done through back propagation, going back from the output layer to the hidden layers. Each cycle of this corresponds to a epoch of our training.



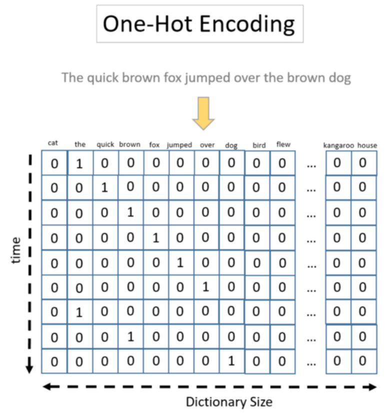
Adjusting the weights can be done using stochastic gradient descent (SGD) to find the optimal point of the loss function, applying it to a random group of data (batch) and frequently updating the parameters once per batch in an epoch.



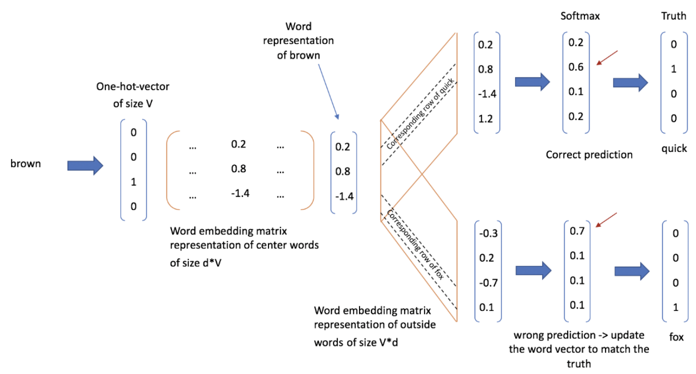
There are different neural network architectures used for different purposes. A recurrent neural network (RNN) is ideal for sentences and other sequence based data, such as time series. RNNs are known to be used in embedding models for text, creating contextual models for sentences, usually through long-term term shot memory (LSTM), encode and decode models or mask models such as Google's BERT architecture.



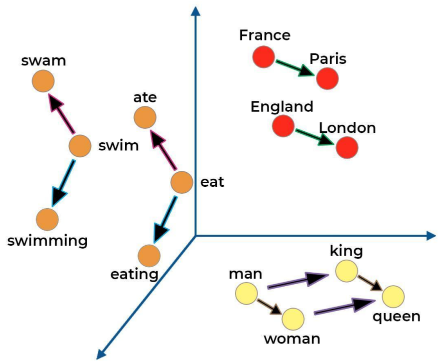
A contextual model creates a binary sparse matrix that considers both the occurrence and position of tokens in the sentence. This huge array resembles the process known as One-Hot Encoding.



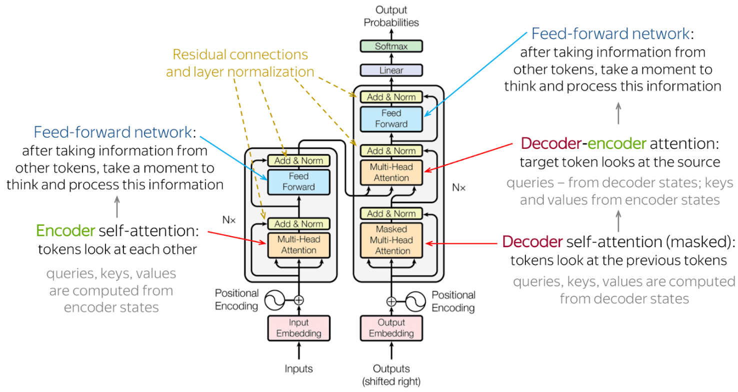
The scalar multiplication of this discrete sparse matrix by continuous random weights is performed within an activation function in neurons contained in the deep layers, transforming it into a dense matrix (less computationally expensive) that is continuous and, therefore, more capable of establishing similarity relations, by distance calculations (e.g. cosine similarity).



Reducing our dimensions and projecting the text embeddings into an orthogonal space, we can, for example, observe the similarity of words from their distance, considering the context of the textual data in which they were trained.



The BERT (Bidirectional Encoder Representations from Transformers) model is a bidirectional mask model that encodes 15% of the sentence data to the left and 15% of the sentence data to the left, passing to the neural networks the ability to predict (decode) this previously known data. The model is extremely complex, having more than 700 dimensions; its architecture (Transformers) is also complex and always ends with a softmax function.

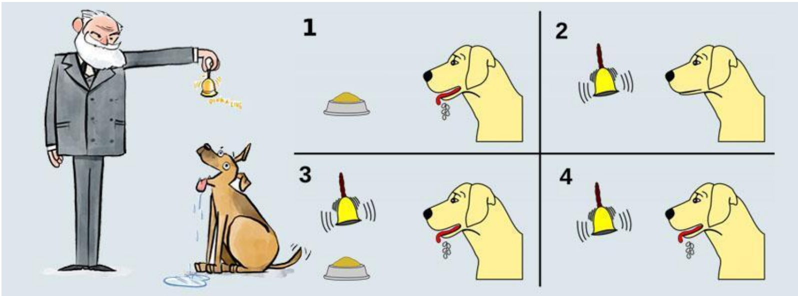


These are just some insights and examples we have about neural networks in machine learning.

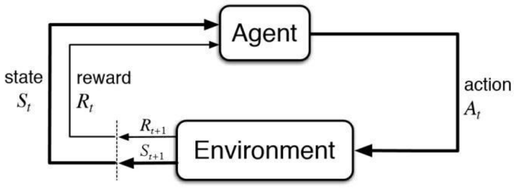
2. *Reinforcement Learning*

I don't have much practical knowledge on the subject. It is a new area that, through self-learning, holds great promise for the automation of various decision-making tasks, such as self-driving cars, marketing and pricing.

I understand it to be comparable to the studies of ethology and behavioral psychology, such as those of Ivan Pavlov. During his research, Pavlov identified an animal behavior in which learning or constancy can be identified or acquired through reward mechanisms, as in humanized training processes (a snack as a reward for each desired behavior). In his experiment, Pavlov taught a dog that whenever a bell rings, he will be fed. Thus, he realized that whenever he rang the bell, the dog already associated it with food and salivated.

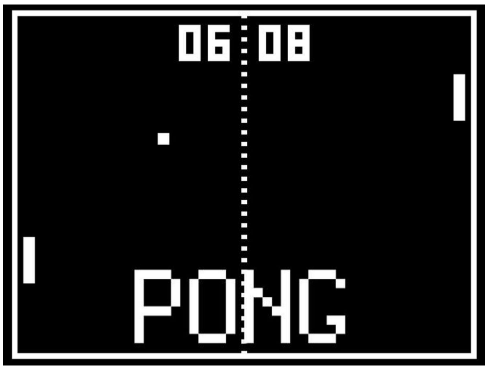


This type of learning has some differences. It’s active, since it presupposes an agent to condition learning. For this, it presupposes a set of rules and objectives (like a game) that can guarantee conditioning through rewards. For this, it also presupposes the existence of a memory. This process repeats itself in a cycle and improves learning.



This has important implications for machine learning and improving our neural networks. Learning requires less data than traditional models, it converges faster and perhaps more accurately. In addition, it is easier for us to use our creativity to direct learning to applications that are in our human interest. It’s much more capable of decision-making for our problems.

Through reinforcement learning, we are able to condition the machine's self-learning for our human or social games. We can teach, for example, the machine to play pong, providing a set of rules (time, displacement) and rewards (points) for an objective (to reach a certain amount of points first).



Reinforcement learning tasks have impressed in recent years. A machine can make chess moves that are unknown to the greatest masters. Recently, AlphaGo has become a self-taught Go player, beating the greatest masters. The product, created by DeepMind, was acquired by Google. It is believed that reinforcement learning could revolutionize the industry by automating the decision-making process for our business problems.