International Journal of Swarm Intelligence and Evolutionary Computation

OPEN ACCESS Freely available online

Editorial

Significance of Deep Learning in Data Mining and Machine Learning

Shorabi P*

Department of Data mining, The University of Iowa, Iowa, USA

DESCRIPTION

Deep learning algorithms use an enormous amount of unsupervised data to automatically extract complex representation. These algorithms are largely motivated by the sector of Artificial Intelligence (AI), which has the overall goal of emulating the human brain's ability to watch, analyze, learn, and make decisions, especially for extremely complex problems. Work concerning these complex challenges has been a key motivation behind Deep Learning algorithms which strive to emulate the hierarchical learning approach of the human brain. Models supported shallow learning architectures like decision trees, support vector machines, and case-based reasoning may come short when attempting to extract useful information from complex structures and relationships in the input corpus. In contrast, Deep Learning architectures have the potential to generalize in non-local and global ways, generating learning patterns and relationships beyond immediate neighbors within the data. Deep learning is actually a crucial step toward AI. It not only provides complex representations of data which are suitable for AI tasks but also makes the machines independent of human knowledge which is the ultimate goal of AI. It extracts representations directly from unsupervised data without human interference.

A key concept underlying Deep Learning methods is distributed representations of the information, during which an outsized number of possible configurations of the abstract features of the input file are feasible, allowing a compact representation of every sample and resulting in a richer generalization. The number of possible configurations is exponentially related to the number of extracted abstract features. Noting that the observed data was generated through interactions of several known/unknown factors, and thus when a knowledge pattern is obtained through some configurations of learnt factors, additional (unseen) data patterns can likely be described through new configurations of the learnt factors and patterns. Compared to learning supported local generalizations, the amount of patterns which will be obtained employing a distributed representation scales quickly with the amount of learnt factors.

Deep learning algorithms cause abstract representations because more abstract representations are often constructed supported less abstract ones. An important advantage of more abstract representations is that they will be invariant to the local changes within the input file. Learning such invariant features is an ongoing major goal in pattern recognition (for example learning features that are invariant to the face orientation in a face recognition task). Beyond being invariant such representations also can disentangle the factors of variation in data. The real data utilized in AI-related tasks mostly arise from complicated interactions of the many sources. For example a picture consists of various sources of variations such a light-weight, object shapes, and object materials. The abstract representations provided by deep learning algorithms can separate the various sources of variations in data.

Stacking up the nonlinear transformation layers are that the basic idea in deep learning algorithms. The more layers the info goes through within the deep architecture, the more complicated the nonlinear transformations which are constructed. These transformations represent the info , so Deep Learning are often considered as special case of representation learning algorithms which learn representations of the info during a Deep Architecture with multiple levels of representations. The achieved final representation may be a highly non-linear function of the input file.

Here we explain two fundamental building blocks, unsupervised single layer learning algorithms which are wont to construct deeper models: Autoencoders and Restricted Boltzmann Machines (RBMs). These are often employed in tandem to construct stacked Autoencoders and Deep belief networks, which are constructed by stacking up Autoencoders and Restricted Boltzmann Machines respectively. Autoencoders, also called autoassociators, are networks constructed of 3 layers: input, hidden and output. Autoencoders attempt to learn some representations of the input within the hidden layer during a way that creates it possible to reconstruct the input within the output layer based on these intermediate representations. Thus, the target output is the input itself. A basic Autoencoder learns its parameters by minimizing the reconstruction error. This minimization is typically done by stochastic gradient

Correspondence to: Shorabi P, Department of Data mining, The University of Iowa, Iowa, USA, E-mail: shorabi@gmail.com

Received date: April 6, 2021; **Accepted date:** April 20, 2021; **Published date:** April 27, 2021

Citation: Shorabi P (2021) Significance of Deep Learning in Data Mining and Machine Learning. Int J Swarm Evol Comput. 10: e215.

descent. If the hidden layer is linear and therefore the mean squared error is employed because the reconstruction criteria, then the Autoencoder will learn the primary k principle components of the info . Alternative strategies are proposed to make Autoencoders nonlinear which are appropriate to build deep networks as well as to extract meaningful representations of data rather than performing just as a dimensionality reduction method.