



Deep Learning Algorithm and Their Applications in the Perception Problem

Redouane Lhiadi

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

January 29, 2020

Deep Learning Algorithm and Their Applications in the Perception Problem

Redouane Lhiadi

CED Science and technology - Faculty
of sciences, University of Mohammed
1st Oujda – Morocco

lhiadi.redouane@gmail.com

ABSTRACT

The objective of this paper is to summarize a comparative account of unsupervised and supervised deep learning models and their applications. The design of a model system requires careful attention to the following issues: definition of pattern classes, sensing environment, pattern representation, feature extraction and selection, cluster analysis, classifier design and learning, selection of training and test samples and performance evaluation. Classification plays a vital role in deep learning algorithms and we found that, though the error backpropagation learning algorithm as provided by supervised learning model is very efficient for a number of non-linear real-time problems, KSOM of unsupervised learning model, offers efficient solution and classification in the perception problem.

Keywords

Deep Learning; Algorithm; Perception; Classification; DL; Learning; Supervised learning; Unsupervised learning.

1. INTRODUCTION

Many ancient machine learning and signal process techniques exploit shallow architectures, that contain one layer of nonlinear feature transformation. samples of shallow architectures square measure typical hidden Andre Mark off models (HMMs), linear or nonlinear resurgent systems, conditional random fields (CRFs), most entropy (MaxEnt) models, support vector machines (SVMs), kernel regression, and multilayer perceptron (MLP) with one hidden layer. A property common to those shallow learning models is that the easy design that consists of just one layer chargeable for remodeling the raw input signals or options into a problem-specific feature area, which can be unperceivable. Take the instance of a support vector machine. it's a shallow linear separation model with one feature transformation layer once kernel trick is employed, and with zero feature transformation layer once kernel trick isn't used. Human scientific discipline mechanisms (e.g., vision and speech), however, recommend the necessity of deep architectures for extracting complicated structure and building cognitive content from made sensory inputs (e.g., natural image and its motion, speech, and music). as an example, human vocalization and perception systems square measure each equipped with clearly bedded ranked structures in remodeling info from the wave shape level to the linguistic level and contrariwise. it's natural to believe that the state of the art may be advanced in process these kinds of media signals if economical and effective deep learning algorithms square measure developed. Signal process systems with deep architectures square measure composed of the many layers of nonlinear process stages, wherever every lower layer's outputs square measure fed to its immediate higher layer because the input. The booming deep learning techniques developed thus far share 2 further key properties: the generative nature of the model, which generally needs a further prime layer to perform the discriminative task, associate degreed an unsupervised pre-training step that

produces effective use of huge amounts of unlabeled coaching knowledge for extracting structures and regularities within the input options.

2. DEFINE LEARNING

Concept learning, which Hunt, Marin, and Stone (1966) describe compactly as "[the] capability to develop classification rules from experience" has long been a principal space of machine learning analysis. supervised idea learning systems are furnished with data concerning many entities whose category membership is thought and turn out from this a characterization of every category. One major dimension on that to differentiate idea learning systems is that the quality of the input and output languages that they use. At one extreme are learning systems that use a propositional attribute-value language for describing entities and classification rules. The simplicity of this formalism permits such systems to influence massive volumes of information and so to take advantage of applied math properties of collections of examples and counterexamples of an idea. At the opposite finish of the spectrum, logical illation systems settle for descriptions of complicated, structured entities and generate classification rules expressed in first-order logic. These usually have access to information pertinent to the domain so need fewer entity descriptions. FOIL, the system delineate during this paper, builds on ideas from each teams. Objects are delineating mistreatment relations and from these FOIL generates classification rules expressed in an exceedingly restricted style of first-order logic, mistreatment strategies custom-made from people who have evolved in attribute-value learning systems [1].

3. A BRIEF HISTORY ABOUT DEEP LEARNING

The thought over deep learning born beside synthetic neural network research Multilayer perceptron with dense stolen layer is a proper example concerning the fashions along flagrant architectures. Backpropagation, invented of 1980s, has been a commonly used algorithm because study the weights concerning these networks. Unfortunately, back-propagation single does no longer employment nicely of action because of discipline networks including greater than a younger number of secret layers (see a decrial then interesting analysis between durability [3]. stability the pervasive appearance over regional optima between the nonconvex objective feature over the sound networks the primary supply about concern of learning. Backpropagation is based on local gradient class then starts generally at incomplete random initial points. It repeatedly gets trapped in poor regional optima or the rapidity will increase extensively as the deepness regarding the networks increases. This subject is partially accountable for steerge away near concerning the computer study or sign technology lookup from neural networks after shallow fashions somuch bear convex deprivation applications (e.g., SVMs, CRFs, then MaxEnt models) because who global greatest be able keep efficiently obtained at the charge regarding less muscular models.

The optimization problem associated together with the deep fashions used to be empirically soothed so a moderately efficient, unsupervised learning algorithm was brought into 2006 by way of Hinton et a durability [2] for a category regarding extreme creative models so those known as awful belief networks (DBNs). A core issue of the DBN is a greedy, layer-by-layer learning algorithm to that amount optimizes DBN weights at time complexity linear in accordance with the quantity then deep over the networks. Separately or with partial surprise, initializing the weights of an MLP including a correspondingly configured DBN often produces a great deal higher effects than so with the random weights [1], [4] As such, deep networks that are learned with unsupervised DBN pretraining followed by the backpropagation fine-tuning are also called DBNs in the literature (e.g.,[5]. and [6]. A DBN comes with additional attractive properties:

- 1) The learning algorithm makes effective use of unlabeled data;
- 2) It can be interpreted as Bayesian probabilistic generative models;
- 3) The values of the hidden variables in the deepest layer are efficient to compute;
- 4) The over fitting problem that is often observed in the models with millions of parameters such as DBNs, and the under fitting problem that occurs often in deep networks are effectively addressed by the generative pre-training step.

The DBN training procedure is not the only one that makes deep learning possible. Since the publication of the seminal work of [7] numerous researchers have been improving and applying the deep learning techniques with success. Another popular technique is to pre-train the deep networks layer by layer by considering each pair of layers as a denoising auto-encoder [2] We will provide a brief overview of the original DBN work and the subsequent progresses in the remainder of this article.

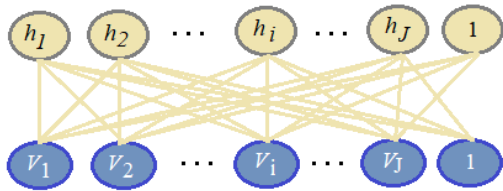


Figure 1. An RBM with i visible units and J hidden units.

4. OVERVIEWS OF DEEP LEARNING AND ITS APPLICATIONS

The primary idea in deep learning algorithm is computerizing the extraction of portrayals (reflections) from the information [8,9,10] Profound learning calculations utilize a tremendous measure of solo information to naturally extricate complex portrayal. These calculations are to a great extent persuaded by the field of man-made consciousness, which has the general objective of imitating the human cerebrum's capacity to watch, break down, learn, and decide, particularly for very perplexing issues. Work relating to these intricate difficulties has been a key inspiration driving Deep Learning calculations which endeavor to imitate the various deep learning approach of the human cerebrum. Models dependent on shallow learning designs, for example, choice trees, bolster vector

machines, and case-based thinking may miss the mark when endeavoring to extricate valuable data from complex structures and connections in the info corpus. Interestingly, Deep Learning structures have the capacity to sum up in non-nearby and worldwide ways, producing learning examples and connections past prompt neighbors in the information [10]

5. DEEP LEARNING APPLICATIONS

Please use a 9-point Times Roman font, or other Roman font with serifs, as close as possible in appearance to Times Roman in which these guidelines have been set. The goal is to have a 9-point text, as you see here. Please use sans-serif or non-proportional fonts only for special purposes, such as distinguishing source code text. If Times Roman is not available, try the font named Computer Modern Roman. On a Macintosh, use the font named Times. Right margins should be justified, not ragged.

5.1 Automatic Speech Recognition (ASR)

The Google has reported that Google voice search had gone in a different direction by receiving Deep Neural Networks (DNN) as the center innovation used to show the hints of a language in 2012 [12] DNN supplanted Gaussian Mixture Model which has been in the business for 30 a long time. DNN additionally has demonstrated that it is better ready to gauge which sound a client is creating at each moment in time and with this they conveyed conspicuously expanded discourse acknowledgment exactness.

In 2013, DL has increased full energy in both ASR and ML [13] DL is fundamentally connected to the utilization of different layers of nonlinear changes to determine discourse highlights, while learning with shallow layers includes the utilization of model based portrayals for discourse highlights which have high dimensionality however normally empty sections.

5.2 Image Recognition

Deep max-pooling convolutional neural systems is utilized to identify mitosis in bosom histology pictures was introduced in [14]. Mitosis location is extremely hard. Truth be told, mitosis is a complex procedure during which a phone core experiences different changes. In this approach, DNN as amazing pixel classifier which works on crude pixel esteems and no human info is required. Subsequently, DNN consequently learns a lot of visual highlights from the preparing information. DNN is tried on a freely accessible dataset and essentially outflanks all contending strategies, with reasonable computational exertion: preparing a 4MPixel picture requires couple of minutes on a standard workstation. Enormous and profound convolutional neural system is prepared to characterize the 1.2 million high resolution pictures in the ImageNet LSVRC-2010 challenge into 1000 distinct classes [15] On the test information, they accomplished top-1 and top-5 mistake paces of 37.5% and 17.0% which is extensively superior to the past cutting edge. From every one of the tests, the outcomes can be improved just by hanging tight for quicker GPUs and greater datasets to wind up accessible.

5.3 Natural Language Processing

As of late, profound learning techniques have been effectively applied to an assortment of language what's more, data recovery applications. By misusing profound structures, profound learning procedures can find from preparing information the concealed structures and highlights at various degrees of deliberations valuable for the any undertakings. In 2013, [16] proposed a progression of Profound Structured Semantic Models (DSSM) for Web search. All the more explicitly, they utilizes a DNN to rank a lot of archives for a given inquiry as pursues. Initial, a non-straight

projection is performed to outline question and the records to a typical semantic space. At that point, the importance of each archive given the inquiry is determined as the cosine similitude between their vectors in that semantic space. The neural system models are discriminatively prepared utilizing the navigate information to such an extent that the restrictive probability of the clicked archive given the inquiry is boosted. The new models are assessed on a Web archive positioning errand utilizing a true informational collection. Results demonstrate that the proposed model fundamentally beats other dormant semantic models, which were viewed as best in class in the execution preceding the work exhibited in [16].

5.4 Drug Discovery and Toxicology

Quantitative Structure Analysis/Prediction Studies (QSAR/QSPR) endeavor to assemble scientific models relating physical and substance properties of mixes to their synthetic structure. In [17], perform multiple tasks learning is applied to QSAR utilizing different neural system models. They utilized a counterfeit neural system to get familiar with a capacity that predicts exercises of mixes for numerous examines simultaneously. The technique is contrasted and elective techniques and revealed that the neural nets with performing multiple tasks can prompt altogether improved outcomes over baselines produced with irregular woods.

In 2015, AtomNet has been presented as first structure-based, profound convolutional neural organize which intended to anticipate the bioactivity of little particles for medication disclosure applications [18] This paper additionally shows how to apply the convolutional ideas of highlight region and progressive organization to the displaying of bioactivity and synthetic collaborations. AtomNet beats past docking approaches on an assorted arrangement of benchmarks by an enormous edge, accomplishing an AUC more noteworthy than 0.9 on 57.8% of the objectives in the DUDE benchmark.

5.5 Customer Relationship Management

A structure for independent control of a client relationship the board framework been diagrammed by [19]. First, an altered form of the generally acknowledged Recency-Frequency Monetary Value arrangement of measurements can be utilized to characterize the state space of customers or benefactors is investigated. Second, a technique to decide the ideal direct showcasing activity in discrete what's more, persistent activity space for the given individual, in view of his situation in the state space is portrayed. The strategy includes the utilization of without model Q-figuring out how to prepare a profound neural arrange that relates a customer's situation in the state space to prizes related with conceivable showcasing exercises. The evaluated worth capacity over the customer state space can be deciphered as client lifetime esteem (CLV), and in this manner takes into consideration a fast module estimation of CLV for a given customer. Test results are introduced, in view of Knowledge Disclosure and Data Mining Tools Competition, mailing dataset of gift sales.

5.6 Recommendation Systems

Programmed music proposal has turned into an undeniably pertinent issue in later a long time, since a great deal of music is currently sold and expended carefully. Most recommender frameworks depend on cooperative separating. In 2013, [20] proposed to utilize an idle factor model for suggestion, and foresee the inert elements from music sound when they can't be acquired from utilization information. Conventional methodology is thought about utilizing a pack of-words portrayal of the sound sign with profound convolutional neural systems, and the expectations is

assessed by quantitatively and subjectively on the Million Song Dataset. The result demonstrates that the ongoing advances in DL make an interpretation of to the music suggestion setting, with profound convolutional neural systems essentially outflanking the conventional approach. Ongoing on the web administrations depend intensely on programmed personalization to prescribe applicable substance to an enormous number of clients. This expects frameworks to scale instantly to oblige the flood of new clients visiting the online administrations just because. Work by [21] in 2015 proposed a substance based suggestion framework to address both the suggestion quality and the framework versatility. They additionally proposed to utilize a rich list of capabilities to speak to clients, as per their web perusing history and search inquiries. They utilize a DL approach to delineate and things to an inactive space where the likeness among clients and their favored things is amplified. Adaptability examination demonstrate that the multi-see DNN model can without much of a stretch scale to envelop a huge number of clients and billions of thing sections.

5.7 Bioinformatics

The comment of genomic data is a noteworthy test in science and bioinformatics. Existing databases of realized quality capacities are inadequate and inclined to mistakes, and the bimolecular trials expected to improve these databases are moderate and exorbitant. While computational strategies are not a substitute for test confirmation, they can help in two different ways: calculations can help in the curation of quality explanations via naturally proposing mistakes, and they can foresee beforehand unidentified quality capacities, quickening the pace of quality capacity revelation. In this work [22], a calculation that accomplishes the two objectives utilizing profound auto encoder neural systems is created. With investigates quality comment information from the Gene Ontology venture, it demonstrates that profound auto encoder systems accomplish preferable execution over other standard AI techniques, including the famous truncated particular worth deterioration.

6. THE RELATIONSHIP BETWEEN LEARNING AND DATA ANALYSIS (LEARNING BASES)

Information examination changes over information into data and learning, and investigates the connection between factors. Information Analysis is the procedure of efficiently applying measurable as well as legitimate methods to depict and outline, consolidate and recap, and assess information. As indicated by Shamoo and Resnik1 different logical methods "give a method for drawing inductive derivations from information and recognizing the sign (the wonder of enthusiasm) from the commotion (factual changes) present in the information." Comprehension of the information examination techniques will empower you to welcome the significance of the logical technique which incorporates testing of speculations and measurable criticalness in connection to research questions. There are various issues that analysts ought to be discerning of as for information examination. A portion of the key contemplations in examination and determination of the correct trial of importance are as per the following:

- Having the fundamental aptitudes to investigate;
- Distinguishing information types;
- Distinguishing various sorts of Statistical tests;
- Identify the determination of a correct test;
- Determining factual essentialness;

- Distinguishing among Parametric and Non-Parametric test with their applying criteria;
- Distinguishing among Correlation and Regression;
- Inappropriate subgroup investigation;
- Lack of unmistakably characterized and target result estimations;
- Partitioning 'content' when investigating subjective information;
- Reliability and Validity;
- Extent of investigation. [23]

Now we can get the relation between the learning bases and the data analysis. And that Develop classification rules from experience" has long been a principal space of machine learning analysis.

7. THE TWO MAIN MODES OF SUPERVISED AND UNSUPERVISED LEARNING, AND THE POSSIBILITY OF COMBINING THEM TO LEARN BETTER

7.1 Introduction

Presentation of intellectual thinking into a traditional PC can take care of issues by model mapping like example acknowledgment, grouping and anticipating. Fake Neural Networks (ANN) gives these sorts of models. These are basically numerical models depicting a capacity; however, they are related with a specific learning calculation or a standard to imitate human activities. ANN is described by three kinds of parameters; (an) in light of its interconnection property (as feed forward system and repetitive system); (b) on its application work (as Classification model, Association model, Optimization model and Self-sorting out model) and (c) in view of the learning rule (managed/solo/support and so forth..) [24]

All these ANN models are remarkable in nature and every offer points of interest of its own. The significant hypothetical and down to earth ramifications of ANN have differing applications. Among these, a great part of the exploration exertion on ANN has concentrated on example order. ANN performs grouping errands clearly and productively due to its auxiliary structure and learning strategies. There is no novel calculation to structure and prepare ANN models since, taking in calculation varies from one another in their learning capacity and level of induction. Consequently, in this paper, we attempt to assess the managed and unaided learning rules and their grouping productivity utilizing explicit model [25]

The general association of the paper is as per the following. After the presentation, we present the different learning calculations utilized in ANN for example arrangement issues and all the more explicitly the learning procedures of directed and unaided calculations in segment II. Area III presents characterization and its necessities in applications and talks about the nature qualification among managed and unaided learning on the example class data. Likewise, we establish framework for the development of characterization arrange for instruction issue of our advantage. Exploratory arrangement and its result of the present investigation are introduced in Section IV. In Section V we examine the final products of these two calculations of the investigation from alternate point of view. Area VI closes with some last musings on regulated and unaided learning calculation for instructive arrangement issue.

7.2 Learning Paradigms

ANN LEARNING PARADIGMS Learning can allude to either gaining or upgrading information. As Herbert Simon says, Machine Learning signifies changes in the framework that are versatile as in they empower the framework to do a similar undertaking or assignments drawn from a similar populace all the more productively and all the more viably whenever. ANN learning ideal models can be delegated administered, unaided and fortification learning. Managed learning model expect the accessibility of an instructor or administrator who characterizes the preparation models into classes and uses the data on the class participation of each preparation occurrence, while, Unsupervised learning model recognize the example class data heuristically and Reinforcement learning learns through experimentation associations with its condition (remunerate/punishment task). In spite of the fact that these models address learning in various ways, learning relies upon the space of interconnection neurons. That is, administered learning learns by changing its entomb association weight mixes with the assistance of blunder signals where as unaided learning utilizes data related with a gathering of neurons and fortification learning utilizes support capacity to adjust nearby weight parameters. Subsequently, learning happens in an ANN by changing the free parameters of the system that are adjusted where the ANN is inserted. This parameter alteration assumes key job in separating the learning calculation as administered or unaided models or different models. Likewise, these learning calculations are encouraged by different learning rules as appeared in the Fig 2 [26].

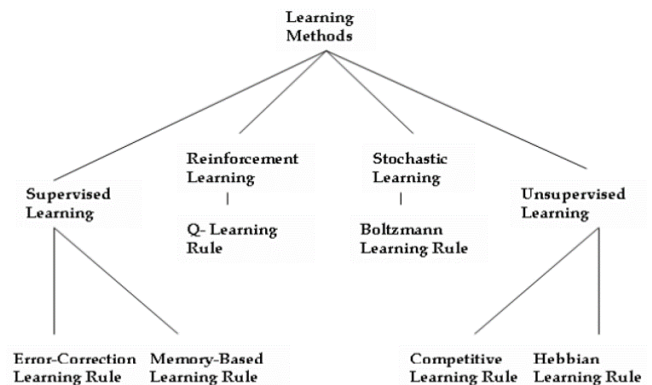


Figure 2 . Learning Rules of ANN.

7.2.1 Supervised Learning

Supervised learning depends on preparing an information test from information source with right order previously doled out. Such strategies are used in feedforward or MultiLayer Perceptron (MLP) models. These MLP has three unmistakable qualities: 1. At least one layers of shrouded neurons that are not part of the info or yield layers of the system that empower the system to learn and tackle any mind-boggling issues 2. The nonlinearity reflected in the neuronal action is differentiable and, 3. The interconnection model of the system shows a high level of network These attributes alongside learning through preparing tackle troublesome and assorted issues. Learning through preparing in an administered ANN model additionally called as mistake backpropagation calculation. The blunder remedy learning calculation prepares the system dependent on the info yield tests and discovers mistake signal, which is the distinction of the yield determined and the ideal yield and alters the synaptic loads of the neurons that is relative to

the result of the mistake signal and the information occurrence of the synaptic weight. In view of this rule, mistake back proliferation learning happens in two passes:

Forward Pass: Here, input vector is exhibited to the system. This info sign spreads forward, neuron by neuron through the system and rises at the yield end of the system as yield signal: $y(n) = \varphi(v(n))$ where $v(n)$ is the prompted neighborhood field of a neuron characterized by $v(n) = \sigma w(n)y(n)$. The yield that is determined at the yield layer $o(n)$ is contrasted and the ideal reaction $d(n)$ and finds the mistake $e(n)$ for that neuron. The synaptic loads of the system during this pass are stays same.

In reverse Pass: The mistake signal that is begun at the yield neuron of that layer is proliferated in reverse through system. This figures the nearby slope for every neuron in each layer and permits the synaptic loads of the system to experience changes as per the delta rule as: $\Delta w(n) = \eta * \delta(n) * y(n)$. This recursive calculation is proceeded, with forward pass pursued by the retrogressive go for each information example till the system is merged [27-28] Regulated learning worldview of an ANN is effective and discovers answers for a few straight and non-direct issues, for example, order, plant control, anticipating, forecast, mechanical technology and so forth [29-30]

Finally, all of two main modes of supervised and unsupervised learning, has its advantages and there is possibility of combining them to learn better.

7.2.2 The Relationship between learning and classification

Classification is one of the most every now and again experienced basic leadership undertakings of human movement. An order issue happens when an item should be doled out into a predefined gathering or class dependent on various watched credits identified with that object. There are numerous modern issues recognized as grouping issues. For models, Stock market expectation, Weather anticipating, Bankruptcy forecast, Medical finding, Speech acknowledgment, Character acknowledgments to give some examples [31-32] These characterization issues can be understood both scientifically and in a non-direct design. The trouble of taking care of such issue numerically lies in the precision and conveyance of information properties and model abilities [33]

The ongoing exploration exercises in ANN demonstrate, ANN as best characterization model due to the non-direct, versatile and useful guess standards. A Neural Network groups a given article as per the yield enactment. In a MLP, when a lot of information examples are exhibited to the system, the hubs in the concealed layers of the system extricate the highlights of the example introduced. For instance, in a 2 concealed layers ANN model, the shrouded hubs in the main shrouded layer structures limits between the example classes and the concealed hubs in the subsequent layer shapes a choice locale of the hyper planes that was framed in the past layer. Presently, the hubs in the yield layer coherently joins the choice locale made by the hubs in the concealed layer and arranges them into class 1 or class 2 as indicated by the quantity of classes portrayed in the preparation with least blunders all things considered. Likewise, in SOM, arrangement occurs by separating highlights by changing of m-dimensional perception info design into q-dimensional element yield space and therefore gathering of articles as per the likeness of the information design.

The motivation behind this investigation is to introduce the reasonable structure of surely understood Supervised and Unsupervised learning calculations in example characterization

situation and to talk about the proficiency of these models in training industry as an example study. Since any order framework looks for a useful connection between the gathering affiliation and property of the item, gathering of understudies in a course for their upgrade can be seen as a characterization issue [34-35] As advanced education has increased expanding significance because of aggressive condition, both the understudies just as the instruction foundations are at junction to assess the presentation and positioning individually. While attempting to hold its high positioning in the training business, every establishment is attempting to distinguish potential understudies and their ranges of abilities and gathering them so as to improve their exhibition and henceforth improve their very own positioning. Thusly, we take this grouping issue and concentrate how the two learning calculations are tending to this issue.

In any ANN model that is utilized for grouping issue, the rule is gaining from perception. As the target of the paper is to watch the example characterization properties of those two calculations, we created Supervised ANN and Unsupervised ANN for the issue referenced previously. A Data set comprises of 10 significant properties that are seen as capability to seek after Master of Computer Applications (MCA), by a college/foundation is taken. These characteristics clarifies, the understudies' scholastic scores, priori science information, score of qualification test directed by the college. Three classes of gatherings are found by the information perception [25].

8. REFERENCES

- [1] J.R. QUINLAN Bassar "Learning Logical Definitions from Relations" Department of Computer Science, University of Sydney, Sydney NSW Australia 2006 Editor: Jack Mostow
- [2] Y. Bengio, "Learning deep architectures for AI," Found. Trends Mach. Learn., vol. 2, no. 1, pp. 1-127, 2009.
G. Hinton, S. Osindero, and Y. Teh, "A fast learning algorithm for deep belief nets," Neural Comput., vol. 18, pp. 1527-1554, 2006
- [3] Learning Logical Definitions from Relations J.R. QUINLAN Bassar Department of Computer Science, University of Sydney, Sydney NSW Australia 2006 Editor: Jack Mostow
- [4] L. Deng, M. Seltzer, D. Yu, A. Acero, A. Mohamed, and G. Hinton, "Binary coding of speech spectrograms using a deep auto-encoder," in Proc. Interspeech, 2010.
- [5] A. Mohamed, D. Yu, and L. Deng, "Investigation of full-sequence training of deep belief networks for speech recognition," in Proc. Interspeech, Sept. 2010.
- [6] A. Mohamed, G. Dahl, and G. Hinton, "Deep belief networks for phone recognition," in Proc. NIPS Workshop Deep Learning for Speech Recognition, 2009.
- [7] G. Hinton, S. Osindero, and Y. Teh, "A fast learning algorithm for deep belief nets," Neural Comput., vol. 18, pp. 1527-1554, 2006.
- [8] Bengio Y, Courville A, Vincent P (2013) Representation learning: A review and new perspectives. Pattern Analysis and Machine Intelligence, IEEE Transactions on 35(8):1798-1828. doi:10.1109/TPAMI.2013.50.
- [9] Bengio Y (2009) Learning Deep Architectures for AI. Now Publishers Inc., Hanover, MA, USA

- [10] Bengio Y (2013) Deep learning of representations: Looking forward. In: Proceedings of the 1st International Conference on Statistical Language and Speech Processing. SLSP'13. Springer, Tarragona, Spain. pp 1–37. http://dx.doi.org/10.1007/978-3-642-39593-2_1.
- [11] Bengio Y, LeCun Y (2007) Scaling learning algorithms towards, AI. In: Bottou L, Chapelle O, DeCoste D, Weston J (eds). Large Scale Kernel Machines. MIT Press, Cambridge, MA Vol. 34. pp 321–360. http://www.iro.umontreal.ca/~lisa/pointeurs/bengio+lecun_chapter2007.pdf
- [12] Hasim Sak, Andrew Senior, Kanishka Rao, Françoise Beaufays and Johan Schalkwyk (September 2015): Google voice search: faster and more accurate.
- [13] Deng, L., & Li, X. (2013). Machine learning paradigms for speech recognition: An overview. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(5), 1060-1089.
- [14] Ciresan, D. C., Giusti, A., Gambardella, L. M., & Schmidhuber, J. (2013, September). Mitosis detection in breast cancer histology images with deep neural networks. In *International Conference on Medical Image Computing and Computer-assisted Intervention* (pp. 411-418). Springer Berlin Heidelberg.
- [15] ger Berlin Heidelberg. [11] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097- 1105).
- [16] Huang, P. S., He, X., Gao, J., Deng, L., Acero, A., & Heck, L. (2013, October). Learning deep structured semantic models for web search using clickthrough data. In *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management* (pp. 2333- 2338). ACM.
- [17] Dahl, G. E., Jaitly, N., & Salakhutdinov, R. (2014). Multi-task neural networks for QSAR predictions. *arXiv preprint arXiv:1406.1231*.
- [18] Wallach, I., Dzamba, M., & Heifets, A. (2015). AtomNet: A Deep Convolutional Neural Network for Bioactivity Prediction in Structure-based Drug Discovery. *arXiv preprint arXiv:1510.02855*.
- [19] Tkachenko, Y. (2015). Autonomous CRM Control via CLV Approximation with Deep Reinforcement Learning in Discrete and Continuous Action Space. *arXiv preprint arXiv:1504.01840*.
- [20] Van den Oord, A., Dieleman, S., & Schrauwen, B. (2013). Deep content-based music recommendation. In *Advances in Neural Information Processing Systems* (pp. 2643-2651).
- [21] Elkahky, A. M., Song, Y., & He, X. (2015, May). A multi-view deep learning approach for cross domain user modeling in recommendation systems. In *Proceedings of the 24th International Conference on World Wide Web* (pp. 278-288). ACM.
- [22] Chicco, D., Sadowski, P., & Baldi, P. (2014, September). Deep autoencoder neural networks for gene ontology annotation predictions. In *Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics* (pp. 533-540). ACM.
- [23] Dr Balkishan Sharma “Processing of data and analysis” Sri Aurobindo Medical College & PG Institute, India
Correspondence: Dr Balkishan Sharma, PhD, Associate Professor (Biostatistics), Department of Community Medicine, Sri Aurobindo, Received: February 06, 2018 | Published: February 20, 2018 Copyright© 2018 Sharma.
- [24] L. Fu., *Neural Networks in Computer Intelligence*, Tata McGraw-Hill, 2003.
- [25] R. Sathya and A. Abraham, “*Application of Kohonen SOM in Prediction*”, In *Proceedings of ICT 2010*, Springer-Verlag Berlin Heidelberg, pp. 313–318, 2010.
- [26] S. Haykin, *Neural Networks- A Comprehensive Foundation*, 2nd ed., Pearson Prentice Hall, 2005.
- [27] R. Rojas, *The Backpropagation Algorithm*, Chapter 7: *Neural Networks*, Springer-Verlag, Berlin, pp. 151-184, 1996.
- [28] Jin-Song Pei, E. Mai, and K. Piyawat, “*Multilayer Feedforward Neural Network Initialization Methodology for Modeling Nonlinear Restoring Forces and Beyond*”, 4th *World Conference on Structural Control and Monitoring*, 2006, pp. 1-8.
- [29] Awodele, and O. Jegede, “*Neural Networks and Its Application in Engineering*”, *Proceedings of Informing Science & IT Education Conference (InSITE) 2009*, pp. 83-95.
- [30] Z. Rao, and F. Alvarruiz, “*Use of an Artificial Neural Network to Capture the Domain Knowledge of a Conventional Hydraulic Simulation Model*”, *Journal of HydroInformatics*, 2007, pg.no 15-24.
- [31] Moghadassi, F. Parvizian, and S. Hosseini, “*A New Approach Based on Artificial Neural Networks for Prediction of High Pressure Vapor-liquid Equilibrium*”, *Australian Journal of Basic and Applied Sciences*, Vol. 3, No. 3, pp. 1851-1862, 2009.
- [32] U. Khan, T. K. Bandopadhyaya, and S. Sharma, “*Classification of Stocks Using Self Organizing Map*”, *International Journal of Soft Computing Applications*, Issue 4, 2009, pp.19-24.
- [33] S. Ali, and K. A. Smith, “*On learning algorithm selection for classification*”, *Applied Soft Computing* Vol. 6, pp. 119–138, 2006.
- [34] L. Schwartz, K. Stowe, and P. Sendall, “*Understanding the Factors that Contribute to Graduate Student Success: A Study of Wingate University’s MBA Program*”.
- [35] V.O. Oladokun, A.T. Adebajo, and O.E. Charles-Owaba, “*Predicting Students’ Academic Performance using Artificial Neural Network: A Case Study of an Engineering Course*”, *The Pacific Journal of Science and Technology*, Vol. 9, pp. 72–79, 2008.