

The upsurge of deep learning for computer vision applications

Priyanka Patel¹, Amit Thakkar²

¹U & P U Patel Department of Computer Engineering, Chandubhai S Patel Institute of Technology,
Faculty of Technology & Engineering, CHARUSAT University, India

²Smt.Kundanben Dinsha Patel Department of Information Technology, Chandubhai S Patel Institute of Technology,
Faculty of Technology & Engineering, CHARUSAT University, India

Article Info

Article history:

Received Feb 26, 2019

Revised Aug 5, 2019

Accepted Aug 29, 2019

Keywords:

Algorithms
Applications
Artificial intelligence
Deep learning
History
Machine learning
Models
Neural network
Neural network rules
Revolution

ABSTRACT

Artificial intelligence (AI) is additionally serving to a brand new breed of corporations disrupt industries from restorative examination to horticulture. Computers can't nevertheless replace humans, however, they will work superbly taking care of the everyday tangle of our lives. The era is reconstructing big business and has been on the rise in recent years which has grounded with the success of deep learning (DL). Cyber-security, Auto and health industry are three industries innovating with AI and DL technologies and also Banking, retail, finance, robotics, manufacturing. The healthcare industry is one of the earliest adopters of AI and DL. DL accomplishing exceptional dimensions levels of accurateness to the point where DL algorithms can outperform humans at classifying videos & images. The major drivers that caused the breakthrough of deep neural networks are the provision of giant amounts of coaching information, powerful machine infrastructure, and advances in academia. DL is heavily employed in each academe to review intelligence and within the trade-in building intelligent systems to help humans in varied tasks. Thereby DL systems begin to crush not solely classical ways, but additionally, human benchmarks in numerous tasks like image classification, action detection, natural language processing, signal process, and linguistic communication process.

Copyright © 2020 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Priyanka Patel,
U & P U Patel Department of Computer Engineering, Chandubhai S Patel Institute of Technology,
CHARUSAT University,
Changa, 388421, Gujarat, India.
Email: priyankapatel.it@charusat.ac.in

1. INTRODUCTION

An interesting aspect regarding neural systems is to what extent they have taken to be an overnight achievement. Andrey Kurenkove illustrate a very nice history chart in Figure 1. History returns the whole distance to the 1940s. Deep learning has extremely just taken off in the last five years from the journey of Neural Network. The reason is the expanded accessibility of name information alongside the significantly expanded computational throughput of current processors.

Deep learning algorithms are composed of algorithms which allows software to instruct and train itself to perform tasks adeptly with the huge dataset of content or text, images, sounds, speeches, videos or time series frames by introducing with multi-layered neural networks with huge amounts of data. Deep learning (DL) is also known like it is the subset of machine learning (ML) made out of calculations that allow programming to prepare itself to perform everyday jobs, similar to discourse and picture acknowledgment, by uncovering multi-layered neural systems to tremendous measures of information.

Artificial Intelligence (AI) could be a machine that solves your real-life problems while not obtaining any humans concerned in it, or to be specific we are able to say that a package that is sort of a little kid that wants some instruction and training and afterward it's capable of resolution your real-life issues. An

assessment of the impact of AI requires a comprehension of both the practice of innovation technology and law. While the referenced examinations assessed exercises of attorneys and current innovations, they don't adequately connect with future conditions of innovation and if you would like a particular definition by Alan Turing, AI defined as "the science of making computers do things that require intelligence when done by humans. Now, Artificial Intelligence could be a big tree that has several branches to review and specialize with it" [1, 2].

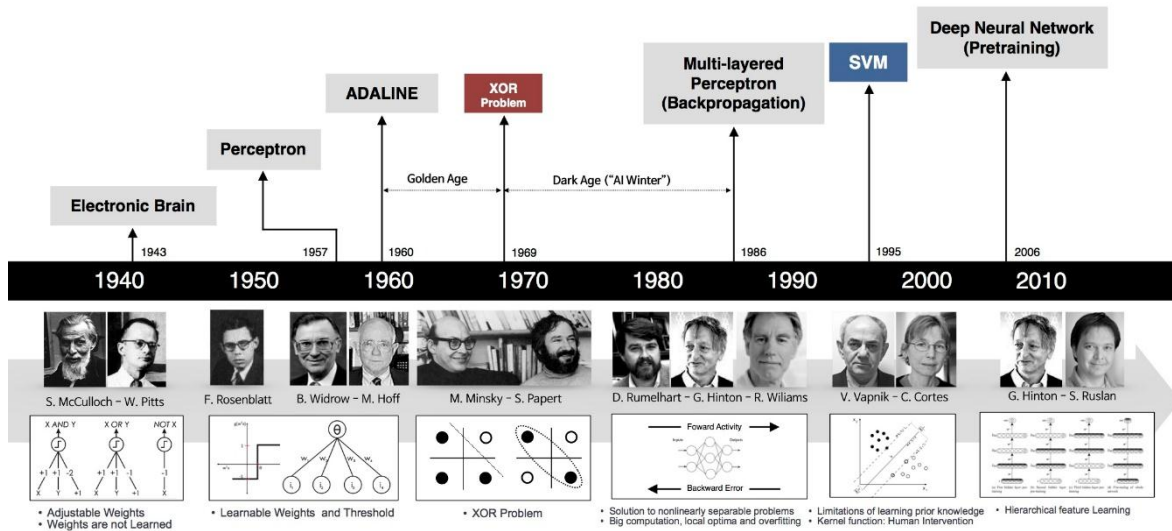


Figure 1. Brief history of neural network and deep learning

Figure 2 is brief of Artificial Intelligence tree and difference between artificial intelligence, Machine Learning, Neural Network, Natural Language Processing, and Deep Learning:

AI: Building systems Building frameworks that can do savvy things implies PCs with the capacity to reason like humans.

ML: It is likewise a subcategory of Artificial Intelligence. Structure a frameworks which will gain as a matter of fact implies machine with the capacity to learn without being unequivocally customized.

NLP: It is a subcategory of Artificial Intelligence. Structure a frameworks that can comprehend language. It is a subcategory of AI.

DL: It is a subcategory of Machine Learning. Structure frameworks that utilization DNN on a substantial arrangement of information implies computers with the ability to learn by using artificial neural networks, which were roused by the model and capacity of the model cerebrum as shwon in Figure 2.

NN: A biologically enlivened network of Artificial Neurons. Artificial Intelligence is encompassed with numerous diverse sub-specialties. Extensively, they can be gathered: Reasoning tools-to generate conclusion from available data or knowledge like picture, content, videos. Unsupervised deep learning tools- create general systems that can be trained with tiny amount of data. The main objective of this kind of learning research is to pre-train a model like discriminator or encoder network to be used for other tasks. Supervised tools-To understand comprehend importance and setting through classification and regression. Need to guide to teach the algorithm what conclusions it should come up with. Learning tools- To put on and broaden understanding through machine learning and predictive analysis. Optimization tools- To put on information to express inquiries through expert frameworks.

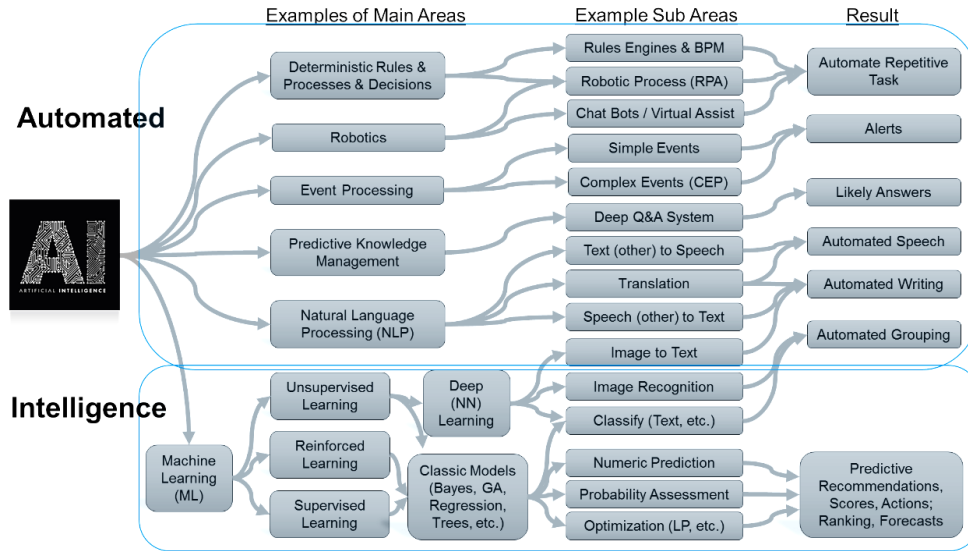


Figure 2. Type of learning and its applications structure systems that use dnn on a large set of data [3, 4]

2. MACHINE LEARNING APPROACH

ML is a kind of artificial brainpower which allows software applications to conclusion with progressively precise in expecting results without being expressly revised. The essential start of ML is to construct calculations which will get input data and utilize measurable examination to predict associate output value within an adequate range. Figure 3 shows the straightforward process of machine learning. In machine learning features are manually extracted and the model creates after extraction of features. These approach follows shallow network as shown in Figure 4. We can improve network through providing more examples and more training data to the network. In deep learning, the features from the dataset are automatically extracted. The Performs of the model is “end-to-end learning” means deep network, not like a shallow network. We can improve the network by providing huger dataset so, a deep learning algorithm is a scale with data. Figure 5 is the example of DL.

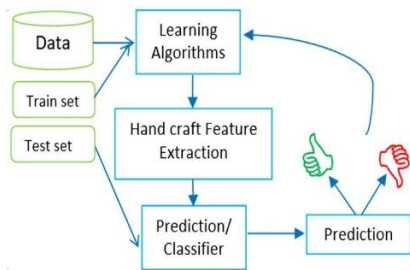


Figure 3. Basic machine learning process flow

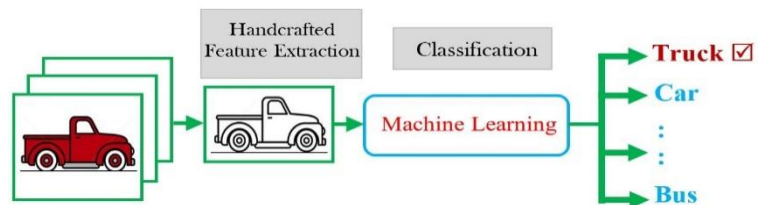


Figure 4. Machine learning approach to hand craft features and classification

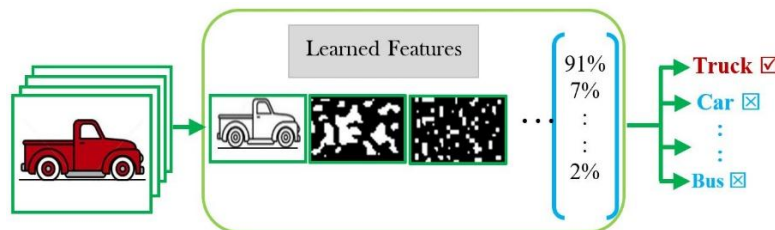


Figure 5. Deep learning approach to auto extracting features and classification [5]

3. NEURAL NETWORKS (NN) AND DEEP LEARNING

NN and DL presently give the most effective solutions to several issues in image recognition, speech recognition, text recognition, time series problems, video processing, and natural language process. Deep learning is nothing but an algorithm which is way smarter than shallow machine learning algorithms but it is not an easy task. Deep neural networks are harder to train and required graphical processing support for better and fast results. Nowadays deep learning applications are used everywhere many examples of deep learning are available in current times such as Google's self-driven car, apples face recognition system, SIRI, Cortana personal assistant all those are developed using deep learning algorithm also newly introduced Amazon Go stores also includes deep learning. Three Capabilities of Deep Learning first is Generalizability—how precisely the machine will execute estimation on given data which has not been formulated yet? Second is Trainability how rapidly decilitre a DL framework will get accustomed to its concern. And third is Expressibility —this feature delineates how well a machine will evaluate general estimations. Deep learning, for instance, interpretability, measured quality, transferability, inertness, ill-disposed soundness, and security.

4. CURRENT CHALLENGES

In current time Deep Learning has turned out to be one in everything about first investigation regions in creating astute machines. The majority of notable applications like Image and Speech Recognition, Text processing, and NPL-Natural language processing of AI are driven by Deep Learning nowadays very easily. Deep Learning algorithms copy human cerebrums exploitation artificial neural networks and more and increasingly more profoundly figure out how to precisely take care of a given issue. However there are vital challenges in Deep Learning systems that we've got to seem out for lots and lots of information, a large amount of data and plenty of information, Overfitting in neural networks, Hyperparameter Optimization, Requires high-performance hardware, Neural networks are in essence a Black-box, Lack of Flexibility and Multitasking.

Deep learning is a methodology that models human theoretical reasoning or possibly speaks to endeavor to approach it rather utilizing. Though, this innovation has an also some set of disadvantages too. Management of Constant Input Data: In DL, a training process depends on examining a lot of information. Albeit, speedily flowing input data and give a little period of time toward guaranteeing a productive training process. Which is the main cause that data researchers need to adjust their algorithm in the manner in which neural networks will deal with a lot of constant input data? Guaranteeing Conclusion Transparency: One more critical disadvantage of deep learning programming is, DL is unequipped for furnished opinions because it has already achieved a convincing conclusion. Unlike in case of ancient machine learning, we can't track an associate algorithm formula to search why, our system has fixed that it's a cat on an image, not a tiger. To address blunders in deep learning program or algorithm, we need to update the entire calculation and algorithm too. Resource-Demanding Technology: It is an altogether resource-demanding technology. It needs additional powerful high-performance GPUs, means its work on CPUs but it gives tremendous results on GPUs, it also required large amounts of storage space to training a model, etc. Moreover, the technology takes a longer training time as compared to ancient ML. Despite all its challenges, deep learning explorers advanced techniques for those researchers who have the intention to use unstructured big data analytics. Indeed, DL tasks of data processing give significant benefits to Educators, analyst, corporations.

5. RECENT DEVELOPMENTS

There are so many computer vision problems and issues like classification reorganization, identification, language processing, video processing, gesture detection, robotics etc. will currently in the world of progressing deep learning area all Computer vision problems be measured as resolved. In recent model-based development, models based on CNN- Convolutional Neural Networks have revolutionized the entire field of computer vision. It is now very easy to develop through several Deep Learning configurations utilizing adjusting with pre-trained weights. For example, classification through ImageNet. Nevertheless, the more security issues of object detection and segmentation need advanced ways to crack the solution. Object detection consists of learning the objects and drawing a rectangle bounding box, whereas segmentation targets to spot precise pixels that fit every object. One amongst most variations through image classification is that a similar image could contain many objects and people might remain in totally diverse sizes, lighting, proportion, and partly occluded. Professor Nick Reed, academy director at the Transport Research Laboratory (TRL), agrees that deep learning is a very important tool, but one that raises serious concerns. "Deep learning is somewhat opaque. It can be hard to understand the rules or knowledge learned by the system. In the case of self-driving cars, this may become important in the event of a collision. For

example, if a vehicle is dependent on a deep learning algorithm, it may be difficult to understand how the vehicle used the available information to determine its actions, subsequently resulting in a lack of clarity over liability."

6. TOWARDS DEEP LEARNING

Deep learning a paradigm of Machine learning which has shown incredible promises in recent time. This is because of the fact that Deep Learning shows a great analogy with the functioning of the human brain. In below Figure 6 shows the Traditional Machine Learning approaches. It seems in a picture of machine learning that features extraction and classification are done separately and both concerned with the complex design and lots of significant mathematics, after doing it very precisely sometimes it was not even efficient, and didn't perform well means in concern with real-world applications the accuracy level wasn't suitable.



Figure 6. Traditional machine learning flow

7. AT PRESENT, ABOUT DEEP LEARNING

DL networks will perform feature extraction and classification in one shot, that's why sometimes people use to call one short learning instead of deep learning. Which implies it just need to plan and design one model. The fundamental advantage of deep learning over ML algorithm is, DL has the ability to create new features from partial series of features located in the training dataset. Also as seen in Figure 7 Feature Extraction and classification will be done in one short. DL model has the convenience of GPUs and GPU will process huge amounts of labeled data in parallel at high speeds empowers. DL models to be much faster than ML methods. In DL Model with backpropagation algorithm, effective loss function like ReLu, and a big amount of parameters, these DL networks are capable to learn extremely complex and abstract features from the set, traditional ML need to handcraft the features and no abstract features are there. Some Pytorch, TensorFlow, and Keras are high-level open source libraries are there with DL frameworks for easy implementation. Nowadays the newest trend in research and development in Machine Learning is Deep learning, why because DL methods have carried ground-breaking advances in computer vision and machine learning. Deep Learning is open, outperform, state-of-art technology. In recent time many new application like speech, video, robotics, security makes practically feasible.

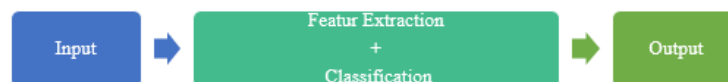


Figure 7. Deep learning flow

8. TYPE OF ARCHITECTURES

Researchers have created a variety of architectures which we can see in below chart that those existed from the late 1940s, architectures and approaches are modified in day by day development and new architectures are created but after 2006 the new architectures and graphical processing units (GPUs) brought them to the pole of AI. From the last two decades, deep learning architectures have greatly expanded the problems neural networks can address. Deep learning is much more than the neural network follows the Figure 8 introduce from many research papers by Favio Vázquez.

Table 1 shows the Neural network to Deep learning models, rules and some functions with their applications. McCulloch–Pitts model, Hebb Rule, Perceptron, Adaptive linear element, Multi-Layer Perceptron (MLP), Hopfield Network Circuit, Back-Propagation, Support vector Machine (SVM), Boltzmann Machine, Restricted Boltzmann Machine (RBM), Deep Boltzmann Machine (DBM), Auto encoder, Recurrent Neural network (RNN), Convolution neural Network or CNN or ConvNet, Long short-term memory (LSTM), Deep belief network (DMN), Deep Auto-encoder, Sparse Auto encoder, Attention based LSTM (ALSTM), Generative Adversarial Network (GAN).

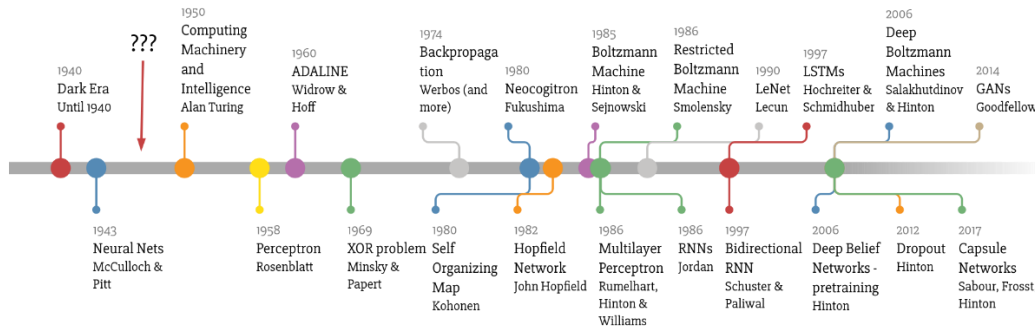


Figure 8. Artificial intelligent, machine learning, deep learning approach's revolution time line

Table 1. Models and rule with their graphical representation and applications

Proposed models	Graphical representation	Applications	Timeline
McCulloch–Pitts model		<ul style="list-style-type: none"> - First ever mathematical model - Linear threshold gate - Weights and bias are fixed - Generates binary output 	1943 [6]
Hebb Rule		<ul style="list-style-type: none"> - It is a learning Rule. - It is used to specify the weight in proportion to activation function. - It works well if the input patterns are orthogonal or uncorrelated. 	1949 [7]
Perceptron		<ul style="list-style-type: none"> - Used for supervised learning of binary classifiers - Classification and Regression 	1958 [8]
Adaptive linear element		<ul style="list-style-type: none"> - Single layer ANN - Adaptive linear unit /neuron 	1960 [9, 10]
Multi-Layer perceptron		<ul style="list-style-type: none"> - Feedforward artificial neural network. - Speech Recognition - Image Recognition - Machine Translation - Classification 	1969 [11, 12]

Table 1. Models and rule with their graphical representation and applications (continue)

Proposed models	Graphical representation	Applications	Timeline
Hopfield Network Circuit		<ul style="list-style-type: none"> - Recurrent Artificial Neural Network - Recognition 	1982 [13]
Back-Propagation	<p>$\sigma(x)$ is the sigmoid function</p> $\sigma(x) = \frac{1}{1 + e^{-x}}$	<ul style="list-style-type: none"> - Classification - Time-series Prediction - Function Approximation 	1970 [14]
Support vector Machine		<ul style="list-style-type: none"> - Bioinformatics - Generalized Predictive Control - Object Detection - Handwriting Recognition - Text And Hypertext Categorization - Images Classification - Protein Fold And Remote Homology Detection 	1963 [15] 1992 [16]
Boltzman Machine		<ul style="list-style-type: none"> - Pattern recognition - Combinatorial optimization - Speech recognition 	1986 [17]
Restricted Boltzman Machine & Deep Boltzman Machine		<ul style="list-style-type: none"> - Pattern Recognition - Classification - Combinatorial Optimization - Speech Recognition 	1986 [18]
Auto encoder		<ul style="list-style-type: none"> - Used for unsupervised learning and dimensionality reduction - It's for representation learning and deep learning - Image generation - Data visualization - Natural language processing 	1986 [19]

Table 1. Models and rule with their graphical representation and applications (*continue*)

Proposed models	Graphical representation	Applications	Timeline
Recurrent Neural network		<ul style="list-style-type: none"> - Speech Recognition - Handwriting Recognition 	1986 [20-23]
Convolution neural Network or CNN or ConvNet		<ul style="list-style-type: none"> - Image And Video Recognition - Image Classification - Video Analysis - Natural Language Processing(NLP) - Medical Image Analysis 	1983 [24] 1999 [25]
Long short-term memory		<ul style="list-style-type: none"> - Speech Recognition - Gesture Recognition - Handwriting Recognition - NLP &Text Compression - Image Captioning 	1997 [26]
Deep belief network		<ul style="list-style-type: none"> - Natural Language Understanding - Image Recognition - Failure Prediction - Information Retrieval 	2007 2009 [27]
Deep Auto encoder		<ul style="list-style-type: none"> - Used for unsupervised learning and dimensionality reduction - It's for representation learning and deep learning - Image generation - Data visualization - Natural language processing 	2000s [28]

Table 1. Models and rule with their graphical representation and applications (continue)

Proposed models	Graphical representation	Applications	Timeline
Sparse Auto encoder	<p>The diagram illustrates a Sparse Autoencoder (SAE) architecture. It starts with an input layer of raw pixels of size $(d_r \times \omega \times r)$. These are processed through three layers: an input layer, a first hidden layer with S_1 units, and a second hidden layer with S_2 units. Weights $W^{(1)}$, $W^{(2)}$, and $W^{(3)}$ connect the layers. The output layer produces reconstructed images and a softmax classifier output S. A graph shows the sigmoid function $\sigma(x)$ used for activation. Labels include 'Input raw pixels of training patches', 'Feature representation with SAE', and 'Softmax classifier training'.</p>	<ul style="list-style-type: none"> - Used for unsupervised learning and dimensionality reduction - Object Recognition - Medical Diagnosis 	2000s [29-31]
Attention based LSTM	<p>The diagram shows an attention-based LSTM. A sequence of features $x_{t-4}, x_{t-3}, x_{t-2}, x_{t-1}, x_t$ is processed by an LSTM hidden layer, producing hidden states h_1, h_2, h_3, h_4, h_5. Attention weights $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ are calculated for each state. These are summed (Σ) and multiplied by a context vector c to produce an attention mechanism output. This output is then processed by a V/A block to produce the final LSTM output.</p>	<ul style="list-style-type: none"> - Used to predict target word - Deep NLP 	2010s [32-34]
Generative Adversarial Network	<p>The diagram depicts a Generative Adversarial Network (GAN). A latent random variable is fed into a generator, which produces synthetic images. These are compared against real world images by a discriminator. The discriminator outputs 'Real' or 'Fake' labels, which are used to calculate a loss function.</p>	<ul style="list-style-type: none"> - unsupervised machine learning - Reconstruct 3D models from image. 	2014 [35-37]

9. HOW DEEP LEARNING WORKS

Deep learning working and the neural network working is almost the same. DL works for both supervised learning and unsupervised learning. And as earlier discussed in paper DL can solve many complex problems of computer vision which were not easily solved by Machine learning. It made up of two principal stages: training and Testing.

- a) The training stages - labeling process of large amounts of data, and define the identical features. The training model compares these features and keeps them to make precise reasoning and conclusion when it encounters related data next time. The DL training process encompasses the following stages:
 - 1) ANNs explore a set of binary false/true questions or.
 - 2) Extract numeral values from data bar.
 - 3) Classify data correspond to the proper responses got.
 - 4) Labelling Data.
- b) The testing stages - label new unexposed data using their previous knowledge and then make a conclusion. In ancient machine learning, new feature extraction will be extracted from a fundamental data stacked into the machine. Analyst formulates ML instructions and corrects the errors encountered by a machine.

This methodology wipes out a negative overtraining impact much of the time be seen in deep learning. In ML, the analyst supplies both examples and training data to assist the system to make correct decisions by the machine this norm is known as supervised learning. In other words, in an ancient machine learning, a computer solves a huge number of tasks, however, it can't mount such undertakings without human control. Diversity between machine learning and deep learning:

- 1) Deep learning requires unlabelled training data in huge quantity to make concise conclusions while Machine Learning will utilize a small amount of data from the analyst.
- 2) Machine learning expects features to be precisely recognized by analyst while deep learning generates new features independently.
- 3) Unlike machine learning, deep learning needs GPU-high-performance hardware.

- 4) Machine learning isolates undertakings into little parts and afterward join and conclude into one output while deep learning solves the problem on the end-to-end basis.
- 5) In comparison with Machine learning, deep learning needs considerably more time to train.
- 6) Machine learning gives enough transparency for its decisions than deep learning.

The idea of deep learning infers that the machine makes its usefulness independent decisions from anyone else as long as it is conceivable at the present time. To surmise, applications of deep learning utilize a various levelled methodology including deciding the most imperative attributes to look at.

10. CONCLUSION

An observations of this paper is Unsupervised learning had been a major effect on in reviving deep learning, but the supervised learning has great success in the 1960s to 2000s so, deep learning has since been overshadowed, but then again after 2000s deep learning is on top to solve supervised and unsupervised complex problems. Although we have not focused to review these kinds of learning. The future expectation for deep learning to become far more important in the longer term. Speech recognition, natural language processing, and video analysis are largely unsupervised: we reviewed the different models of deep learning to solve these kinds of problems. Deep learning algorithms and models are trained end-to-end and that uses reinforcement learning to decide where to look. Deep learning and reinforcement learning already outperform in one short classification and feature extraction and produce impressive results. Natural language processing is another area in which deep learning is doing large impact of experiments over the next few years. Speech recognition and video analysis are also major complex issues in computer vision. For example NLP problem, we expect a model which use Recurrent Neural Network to understand documents content or sentences will become the better rule for selectively attending to one part at a time. Eventually, major progress in artificial intelligence will come about through systems deep learning. In spite of the fact that to produce good result deep learning has been used for image, speech, video and handwriting recognition.

REFERENCES

- [1] H. A. Simon, "Artificial intelligence: an empirical science," *Artificial Intelligence*, vol. 77, pp. 95-127, 1995.
- [2] T. Dettmers, T. Dettmers, T. Dettmers, E. Shelhamer and T. Dettmers, "Deep Learning in a Nutshell: History and Training | NVIDIA Developer Blog", NVIDIA Developer Blog, 2019. [Online]. Available: <https://devblogs.nvidia.com/deep-learning-nutshell-history-training>.
- [3] <https://www.xenonstack.com/blog/data-science/overview-of-artificial-intelligence-and-role-of-natural-language-processing-in-big-data>
- [4] https://i2.wp.com/vincejeffs.com/wp-content/uploads/2017/03/AI_Automated_Intelligence.png
- [5] <https://www.mathworks.com/discovery/deep-learning.html>.
- [6] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol. 5, pp. 115-133, 1943.
- [7] D. O. Hebb, "The organization of behavior. A neuropsychological theory," 1949.
- [8] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain," *Psychological review*, vol. 65, pp. 386, 1958.
- [9] B. Widrow, "Adaptive adaline Neuron Using Chemical," *memistors*, 1960.
- [10] Eriksson, K., & Johnson, C. (1988). An adaptive finite element method for linear elliptic problems. *Mathematics of Computation*, 50(182), 361-383.
- [11] M. L. Minsky and S. A. Papert, "Perceptrons: an introduction to computational geometry," MA, MIT Press, Cambridge, 1969.
- [12] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988). Learning representations by back-propagating errors. *Cognitive modeling*, 5(3), 1.
- [13] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the national academy of sciences*, vol. 79, pp. 2554-2558, 1982.
- [14] D. E. Rumelhart, *et al.*, "Learning representations by back-propagating errors," *nature*, vol. 323, pp. 533, 1986.
- [15] Vapnik and Lerner, "Introduce the Generalized Portrait algorithm (the algorithm implemented by support vector machines is a nonlinear generalization of the Generalized Portrait algorithm)," 1963.
- [16] V. Vapnik and A. Lerner, "Pattern recognition using generalized portrait method," *Automation and Remote Control*, vol. 24, pp. 774-780, 1963.
- [17] D. H. Ackley, *et al.*, "A learning algorithm for Boltzmann machines," *Cognitive science*, vol. 9, pp. 147-169, 1985.
- [18] P. Smolensky, "Information processing in dynamical systems: Foundations of harmony theory," Colorado Univ at Boulder Dept of Computer Science, 1986.
- [19] D. E. Rumelhart, *et al.*, "Learning internal representations by error propagation," in Rumelhart D., *et al.*, "Parallel distributed processing: Explorations in the microstructure of cognition," *Foundations*, vol. 1, 1986.
- [20] R. J. Williams, *et al.*, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533-536, 1986.

- [21] K. Fukushima and S. Miyake, "Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition," *Competition and cooperation in neural nets*, Springer, Berlin, Heidelberg, pp. 267-285, 1982.
- [22] Liang, M., & Hu, X. (2015). Recurrent convolutional neural network for object recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3367-3375).
- [23] Lange, S., & Riedmiller, M. (2010, July). Deep auto-encoder neural networks in reinforcement learning. In The 2010 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.
- [24] K. Fukushima, S. Miyake and T. Ito, "Neocognitron: A neural network model for a mechanism of visual pattern recognition", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. -13, no. 5, pp. 826-834, 1983. Available: 10.1109/tsmc.1983.6313076.
- [25] Y. LeCun, *et al.*, "Object recognition with gradient-based learning," *Shape, contour and grouping in computer vision*, Springer, Berlin, Heidelberg, pp. 319-345, 1999.
- [26] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, pp. 1735-1780, 1997.
- [27] G. E. Hinton, "Deep belief networks," *Scholarpedia*, vol. 4, pp. 5947, 2009.
- [28] Y. Bengio and Y. LeCun, "Scaling learning algorithms towards AI," *Large-scale kernel machines*, vol. 34, pp. 1-41, 2007.
- [29] Ng, A. (2011). Sparse autoencoder. CS294A Lecture notes, 72(2011), 1-19.
- [30] C. Doersch, "Tutorial on variational autoencoders," arXiv preprint arXiv:1606.05908, 2016.
- [31] Sun, W., Shao, S., Zhao, R., Yan, R., Zhang, X., & Chen, X. (2016). A sparse auto-encoder-based deep neural network approach for induction motor faults classification. *Measurement*, 89, 171-178.
- [32] Y. Qin, *et al.*, "A dual-stage attention-based recurrent neural network for time series prediction," arXiv preprint arXiv:1704.02971, 2017.
- [33] Wang, Y., Huang, M., Zhu, X., & Zhao, L. (2016, November). Attention-based LSTM for aspect-level sentiment classification. In Proceedings of the 2016 conference on empirical methods in natural language processing (pp. 606-615).
- [34] Gao, L., Guo, Z., Zhang, H., Xu, X., & Shen, H. T. (2017). Video captioning with attention-based LSTM and semantic consistency. *IEEE Transactions on Multimedia*, 19(9), 2045-2055.
- [35] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems* (pp. 2672-2680).
- [36] Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2016). Improved techniques for training gans. In *Advances in neural information processing systems* (pp. 2234-2242).
- [37] Jahanian, A., Chai, L., & Isola, P. (2019). On the "steerability" of generative adversarial networks. arXiv preprint arXiv:1907.07171.

BIOGRAPHIES OF AUTHORS



Priyanka P. Patel has received her Bachelor degree in Information Technology from Gujarat University in the year 2006. She has completed her Master degree in Computer Engineering from Chandubhai S Patel Institute of Technology, CHARUSAT university in the Year 2013. She is currently pursuing her PhD degree in computer engineering at CHARUSAT University, Gujarat, India, and working as assistant professor in Department of Information Technology, Chandubhai S. Patel of Institute, Gujarat, India. Her research interests include image processing, Video Analysis, Machine Learning, Deep Learning.



Dr. Amit Thakkar has received his B.E Degree in I.T. from Gujarat University in 2002 and Master Degree from Dharmsinh Desai University, Gujarat, India in 2007. He has finished his PhD in the area of Multi Relational Classification from Kadi Sarva Vishwa Vidyalyaya (KSV), Gandhinagar, India in 2016. He is working as an Associate Professor in the Department of Information Technology in Faculty of Engineering and Technology, CHARUSAT University, Changa, Gujarat Since 2002. He has published more than 50 research papers in international, national journals and conferences of repute. He has more than fifteen years of teaching experience. He has developed subject proficiency in Data Structures, Database Management System, Data Compression and Data Mining His research interests include database, Data Mining, Machine Learning, Deep Learning.