ISSN: 2456-9348
Vol (02) _Issue (02) Impact Factor: 4.520

iJETRM
International Journal of Engineering Technology Research & Management

OPPORTUNITIES AND GAINSAYS IN DEEP LEARNING ON BIG DATA

Tanya Tiwari*[1],
Tanuj Tiwari[2],
Sanjay Tiwari[3],
[*1]Samsung India Electronics Pvt.Ltd. Noida
[2]Research Assistant, University of Texas at Dallas, Richardson USA
[2]Photonics Res. lab., S.O.S. in Electronics & Photonics, Pt. Ravishankar Shukla University, Raipur(CG),India

ABSTRACT
Deep learning is currently an extremely active research area in pattern recognition society. It has gained huge successes in a broad area of applications such as speech recognition, computer vision, and natural language processing. With the sheer size of data available today, big data brings big opportunities and transformative potential for various sectors; on the other hand, it also presents unprecedented challenges to harnessing data and information. As the data keeps getting bigger, Deep learning is coming to play a key role in providing big data predictive analytics solutions.Big data assist ML algorithms to uncover more fine-grained patterns and helps in accurate predictions .The major challenges to ML are model scalability and distributed computing.The realization of this grand potential relies on the ability to extract value from such massive data through data analytics; Deep learning is at its core because of its ability to learn from data and provide data driven insights, decisions, and predictions. In this paper first, we review the Deep learning techniques and highlight some promising learning methods in recent studies, such as representation learning, deep learning, distributed and parallel learning, transfer learning, active learning, and kernel-based learning and analyse the challenges and possible solutions of Deep learning for big data. Finally, we outline several open issues and research trends.

Keywords:
Artificial Intelligence, DeepLearning, Deep Learning

INTRODUCTION
Deep learning and Big Data are two hottest trends in the rapidly growing digital world. While Big Data has been defined in different ways, herein it is referred to the exponential growth and wide availability of digital data that are difficult or even impossible to be managed and analyzed using conventional software tools and technologies. Digital data, in all shapes and sizes, is growing at astonishing rates. [1-2]. For example, according to the National Security Agency, the Internet is processing 1,826 Petabytes of data per day [3-5]. In 2011, digital information has grown nine times in volume in just five years [6] and by 2020, its amount in the world will reach 35 trillion gigabytes [7-8]. This explosion of digital data brings big opportunities and transformative potential for various sectors such as enterprises, healthcare industry manufacturing, and educational services [4]. It also leads to a dramatic paradigm shift in our scientific research towards data-driven discovery. Traditional approaches to programming rely on hardcoded rules, which set out how to solve a problem, step-by-step. In contrast, Deep learning systems are set a task, and given a large amount of data to use as examples of how this task can be achieved or from which to detect patterns. The system then learns how best to achieve the desired output. It can be thought of as narrow AI: Deep learning supports intelligent systems, which are able to learn a particular function, given a specific set of data to learn from[5-8]. AI is the all-encompassing umbrella that covers everything from Good Old Fashion AI (GOFAI) all the way to connectionist architectures like Deep Learning. ML is a sub-field of AI that covers anything that has to do with the study of learning algorithms by training with data. There are whole swaths (not swatches) of techniques that have been developed over the years like Linear Regression, K-means, Decision Trees, Random Forest, PCA, SVM and finally Artificial Neural Networks (ANN). Artificial Neural Networks is where the field of Deep Learning had its genesis from.
Deep Learning and Big Data as such have no direct relation. Although one can say that Big Data Techniques can be used in Deep Learning.DeepLearning usually works with huge chunks of data and this where Big Data comes into picture.

iJETRM
International Journal of Engineering Technology Research & Management

Deep learning (ML) is continuously unleashing its power in a wide range of applications. It has been pushed to the forefront in recent years partly owing to the advent of big data. ML algorithms have never been better promised while challenged by big data. Big data enables ML algorithms to uncover more fine-grained patterns and make more timely and accurate predictions than ever before; on the other hand, it presents major challenges to ML such as model scalability and distributed computing.

However, traditional Deep learning approaches were developed in a different era, and thus are based upon multiple assumptions, such as the data set fitting entirely into memory, what unfortunately no longer holds true in this new context. These broken assumptions, together with the Big Data characteristics, are creating obstacles for the traditional techniques. Consequently, this paper compiles, summarizes, and organizes Deep learning challenges with Big Data. In contrast to other research that discusses challenges, this work highlights the cause-effect relationship by organizing challenges according to Big Data Vs or dimensions that instigated the issue: volume, velocity, variety, or veracity. Moreover, emerging Deep learning approaches and techniques are discussed in terms of how they are capable of handling the various challenges with the ultimate objective of helping practitioner's select appropriate solutions for their use cases. Finally, a matrix relating the challenges and approaches is presented. Through this process, this paper provides a perspective on the domain, identifies research gaps and opportunities, and provides a strong foundation and encouragement for further research in the field of Deep learning with Big Data.

THE DEEP LEARNING TECHNIQUES

Recent years have seen exciting advances in Deep learning, which have raised its capabilities across a suite of applications. Increasing data availability has allowed Deep learning systems to be trained on a large pool of examples, while increasing computer processing power has supported the analytical capabilities of these systems. Within the field itself there have also been algorithmic advances, which have given Deep learning greater power. As a result of these advances, systems which only a few years ago performed at noticeably below-human levels can now outperform humans at some specific tasks. Many people now interact with systems based on Deep learning every day, for example in image recognition systems, such as those used on social media; voice recognition systems, used by virtual personal assistants; and recommender systems, such as those used by online retailers. As the field develops further, Deep learning shows promise of supporting potentially transformative advances in a range of areas, and the social and economic opportunities which follow are significant. In healthcare, Deep learning is creating systems that can help doctors give more accurate or effective diagnoses for certain conditions. In transport, it is supporting the development of autonomous vehicles, and helping to make existing transport networks more efficient. For public services it has the potential to target support more effectively to those in need, or to tailor services to users. And in science, Deep learning is helping to make sense of the vast amount of data available to researchers today, offering new insights into biology, physics, medicine, the social sciences, and more[1-12].

Deep-learning technology powers many aspects of modern society: from web searches to content filtering on social net- works to recommendations on e-commerce websites, and it is increasingly present in consumer products such as cameras and smartphones.Deep-learning systems are used to identify objects in images, transcribe speech into text, match news items, posts or products with users' interests, and select relevant results of search. Increasingly, these applications make use of a class of techniques called deep learning. Deep learning can be used to extract knowledge from data, learn tasks that are difficult to formalise and create software that improves over time

Conventional Deep-learning techniques were limited in their ability to process natural data in their raw form. For decades, constructing a pattern-recognition or Deep-learning system required careful engineering and considerable domain expertise to design a feature extractor that transformed the raw data (such as the pixel values of an image) into a suitable internal representation or feature vector from which the learning subsystem, often a classifier, could detect or classify patterns in the input.

Generally, the field of Deep learning is divided into three subdomains: supervised learning, unsupervised learning, and reinforcement learning.

Supervised Deep learning

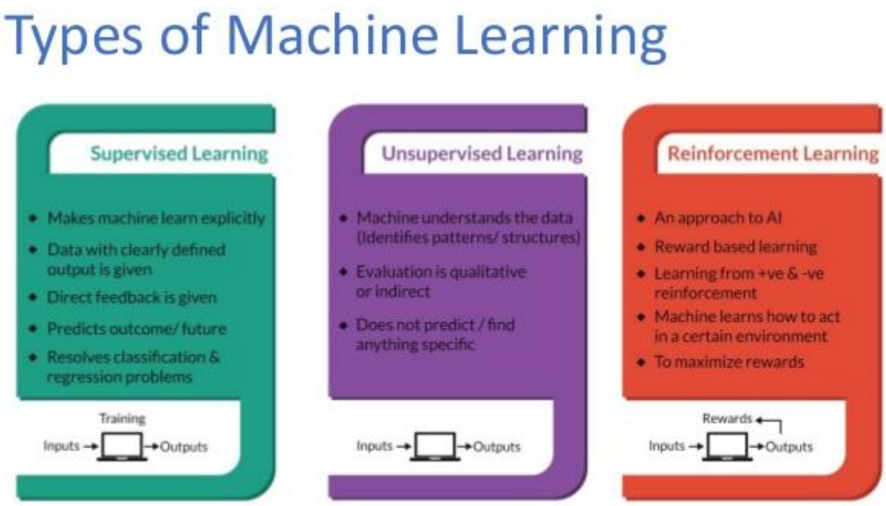
It includes such algorithms as linear and logistic regression, multi-class classification, and support vector Deeps. Supervised learning is so named because the data scientist acts as a guide to teach the algorithm what conclusions it should come up with. It's similar to the way a child might learn arithmetic from a teacher. Supervised learning requires that the algorithm's possible outputs are already known and that the data used to

iJETRM
International Journal of Engineering Technology Research & Management

train the algorithm is already labeled with correct answers. For example, a classification algorithm will learn to identify animals after being trained on a dataset of images that are properly labeled with the species of the animal and some identifying characteristics.

Unsupervised Deep learning

It is more closely aligned with what some call true artificial intelligence — the idea that a computer can learn to identify complex processes and patterns without a human to provide guidance along the way. Although unsupervised learning is prohibitively complex for some simpler enterprise use cases, it opens the doors to solving problems that humans normally would not tackle. Some examples of unsupervised Deep learning algorithms include k-means clustering, principal and independent component analysis, and association rules. While a supervised classification algorithm learns to ascribe inputted labels to images of animals, its unsupervised counterpart will look at inherent similarities between the images and separate them into groups accordingly, assigning its own new label to each group. In a practical example, this type of algorithm is useful for customer segmentation because it will return groups based on parameters that a human may not consider due to pre-existing biases about the company's demographic. The selection of supervised or unsupervised Deep learning algorithm typically depends on factors related to the structure and volume of your data and the use case of the issue at hand. In **Semi-supervised** learning some data is labeled but most of it is unlabeled and a mixture of supervised and unsupervised techniques can be used.



Fig,1 Classification of Deep Learning

iJETRM
International Journal of Engineering Technology Research & Management

Table 1 Comparison of machine learning technologies

Learning types	Data processing tasks	Distinction norm	Learning algorithms
Supervised learning	Classification/Regression/Estimation	Computational classifiers	Support vector machine
		Statistical classifiers	Naïve Bayes
			Hidden Markov model
			Bayesian networks
		Connectionist classifiers	Neural networks
Unsupervised learning	Clustering/Prediction	Parametric	K-means
			Gaussian mixture model
		Nonparametric	Dirichlet process mixture model
			X-means
Reinforcement learning	Decision-making	Model-free	Q-learning
			R-learning
		Model-based	TD learning
			Sarsa learning

Fig.2 Comparision of Deep Learning

ADVANCED LEARNING METHODS

In the following section characteristic of advanced learning methods which are based on learning, rather than just a single algorithm are discussed.

1. **Representation learning** is a set of methods that allows a Deep to be fed with raw data and to automatically discover the representations needed for detection or classification. Deep-learning methods are representation-learning methods with multiple levels of representation, obtained by composing simple but non-linear modules that each transform the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level. With the composition of enough such transformations, very complex functions can be learned. For classification tasks, higher layers of representation amplify aspects of the input that are important for discrimination and suppress irrelevant variations. An image, for example, comes in the form of an array of pixel values, and the learned features in the first layer of representation typically represent the presence or absence of edges at particular orientations and locations in the image. The second layer typically detects motifs by spotting particular arrangements of edges, regardless of small variations in the edge positions. The third layer may assemble motifs into larger combinations that correspond to parts of familiar objects, and subsequent layers would detect objects as combinations of these parts. The key aspect of deep learning is that these layers of features are not designed by human engineers: they are learned from data using a general-purpose learning procedure.
2. **Deep Learning:** Deep Learning, as the term "deep" specifies is inspired by the human brain and it consists of artificial neural networks (ANN) that are modelled on a similar architecture present in the human brain. In Deep Learning, the learning is performed through a deep and multi-layered "network" of interconnected "neurons". The term "deep" usually refers to the number of hidden layers in the neural network. According to a Mathwork blog, traditional neural networks only contain 2-3 hidden layers, while deep networks can have as many as 150. In 2006, Geoffrey Hinton coined the term "deep learning" to explain new algorithms that allow computers to distinguish objects and text in images and videos. Deep-learning theory shows that deep nets have two different exponential advantages over classic learning algorithms that do not use distributed representations. Both of these advantages arise from the power of composition and depend on the underlying data-generating distribution having an appropriate componential structure. First, learning distributed representations enable generalization to new combinations of the values of learned features beyond those seen during training (for example, 2n

iJETRM
International Journal of Engineering Technology Research & Management

combinations are possible with n binary features). Second, composing layers of representation in a deep net brings the potential for another exponential advantage (exponential in the depth).

3. **Distributed and parallel learning**: There is often exciting information hidden in the unprecedented volumes of data. Learning from these massive data is expected to bring significant science and engineering advances which can facilitate the development of more intelligent systems. However, a bottleneck preventing such a big blessing is the inability oflearning algorithms to use all the data to learn within a reasonable time. In this context, distributed learning seems to be a promising research since allocating the learning process among several workstations is a natural way of scaling up learning algorithms. Different from the classical learning framework, in which one requires the collection of that data in a database for central processing, in the framework of distributed learning, the learning is carried out in a distributed manner.
4. **The growing amount of available information and its distributed and heterogeneous nature has a major impact on the field of data mining.**A framework is to be developed for parallel and distributed boosting algorithms intended for efficient integrating specialized classifiers learned over very large, distributed and possibly heterogeneous databases that cannot fit into main computer memory. Boosting is a popular technique for constructing highly accurate classifier ensembles, where the classifiers are trained serially, with the weights on the training instances adaptively set according to the performance of previous classifiers. The parallel boosting algorithm is to be designed for tightly coupled shared memory systems with a small number of processors, with an objective of achieving the maximal prediction accuracy in fewer iterations than boosting on a single processor. After all processors learn classifiers in parallel at each boosting round, they are combined according to the confidence of their prediction. Our distributed boosting algorithm is proposed primarily for learning from several disjoint data sites when the data cannot be merged together, although it can also be used for parallel learning where a massive data set is partitioned into several disjoint subsets for a more efficient analysis
5. **Transfer learning:** A major assumption in many traditional Deep learning algorithms is that the training and test data are drawn from the same feature space and have the same distribution.However, with the data explosion from variety of sources, great heterogeneity of the collected data destroys the hypothesis. To tackle this issue, transfer learning has been proposed to allow the domains, tasks, and distributions to be different, which can extract knowledge from one or more source tasks and apply the knowledge to a target task .The advantage of transfer learning is that it can intelligently apply knowledge learned previously to solve new problems faster
6. **Active learning:** In many real-world applications, we have to face such a situation: data may be abundant but labels are scarce or expensive to obtain. Frequently, learning from massive amounts of unlabeled data is difficult and time-consuming. Active learning attempts to address this issue by selecting a subset of most critical instances for labeling . In this way, the activelearner aims to achieve high accuracy using as few labelled instances as possible, thereby minimizing the cost of obtaining labelled data . It can obtain satisfactory classification performance with fewer labelled samples via query strategies than those of conventional passive learning.
7. **Kernel-based learning:** Over the last decade, kernel-based learning has established itself as a very powerful technique to increase the computational capability based on a breakthrough in the design of efficient nonlinear learning algorithms . The outstanding advantage of kernel methods is their elegant property of implicitly mapping samples from the original space into a potentially infinite-dimensional feature space, in which inner products can be calculated directly via a kernel function .

iJETRM
International Journal of Engineering Technology Research & Management

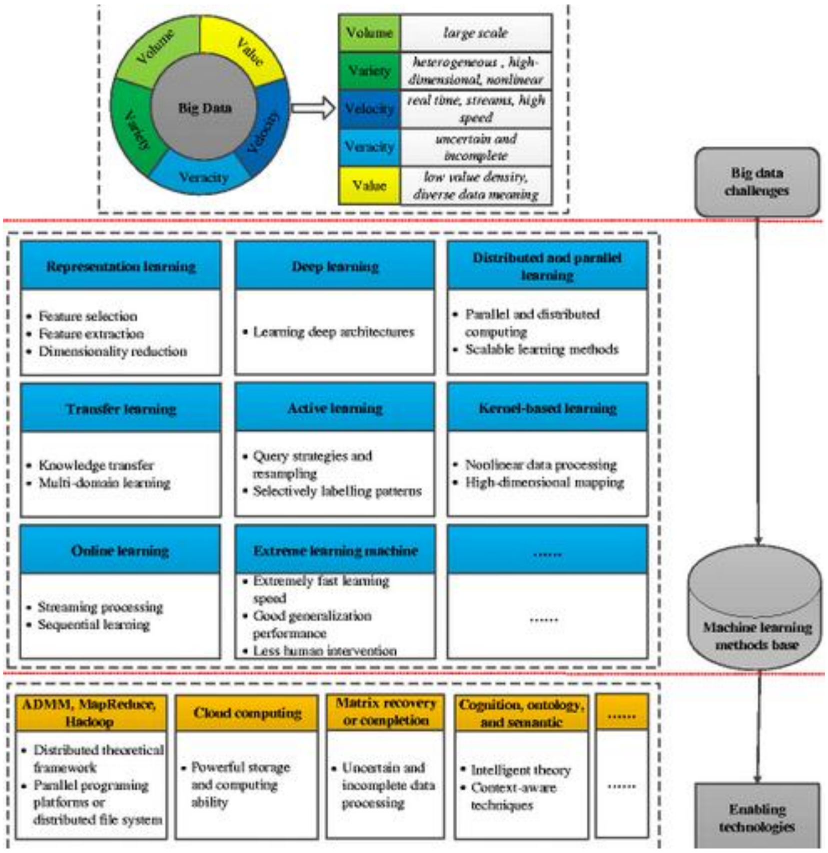


Fig.3 Hierarchical framework of efficient Deep learning for big data processing

Deep learning is just a part of data science. Data science is a big umbrella covering each and every aspect of data processing and not only statistical or algorithmic aspects. To mention, data science includes

- data visualization
- data integration
- dashboards and BI
- distributed architecture
- automated, data-driven decisions
- automating Deep learning
- deployment in production mode
- data engineering

Deep learning helps data science by making a provision for data analysis, data preparation and even decision making like real time testing, online learning. Data science clubs together algorithms derived from Deep learning in order to provide a solution. Data science carries out this activity by taking a lot of ideas from basic mathematics, statistics and domain expertise.

ISSN: 2456-9348
Vol (02) _Issue (02) Impact Factor: 4.520

iJETRM
International Journal of Engineering Technology Research & Management

BIG DATA ANALYTICS

Big Data Analytics is studying large datasets (big data) to identify hidden patterns, market trends, consumer preferences and other valuable information helping organizations to form strategic business decisions.With Big data analytics, data scientists and other analytics professionals can examine huge amounts of structured data as well as the untapped data by deploying analytics and business intelligence.Big Data Analytics comprises of specialized software and analytics systems benefiting business in many ways like

- Cost efficiency: Hadoop and cloud based analytics are big data analytics technologies are very cost effective when storing huge amounts of data. Moreover, this also helps in finding more effectual ways of doing business.
- Faster decision making: Organizations can examine data immediately with superfast Hadoop and in-memory analytics. Decisions can take with much ease on the basis of what they have experienced.
- New products and services: Big data analytics helps to easily understand consumer needs and preferences giving more power to serve customers what they want. More products and services can be developed to fulfill customer's needs.

are important for discrimination and suppress irrelevant variations. An image, for example, comes in the form of an array of pixel values, and the learned features in the first layer of representation typically represent the presence or absence of edges at particular orientations and locations in the image. The second layer typically detects motifs by spotting particular arrangements of edges, regardless of small variations in the edge positions. The third layer may assemble motifs into larger combinations that correspond to parts of familiar objects, and subsequent layers would detect objects as combinations of these parts. The key aspect of deep learning is that these layers of features are not designed by human engineers: they are learned from data using a general-purpose learning procedure.

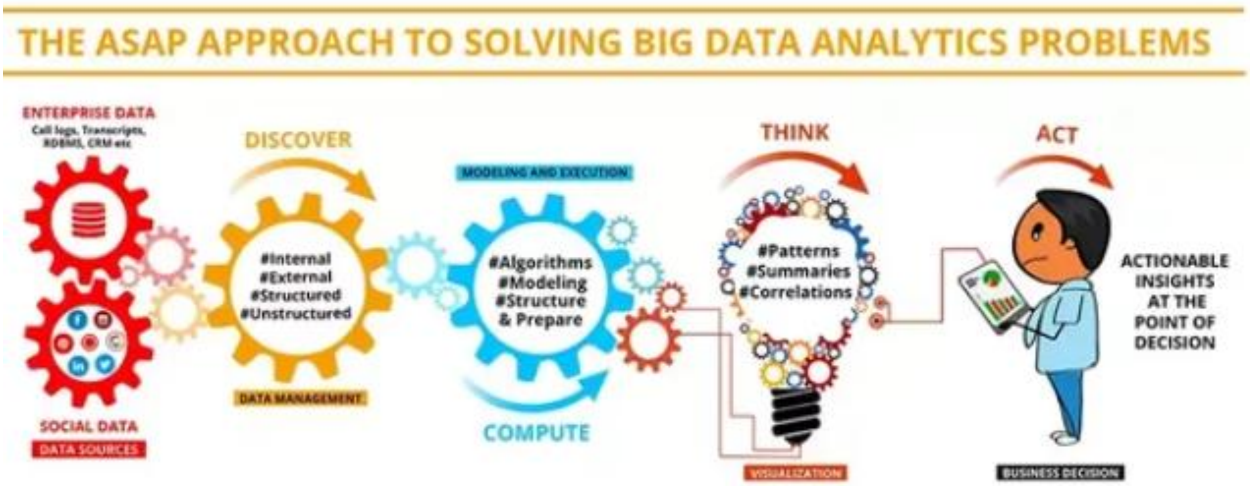


Fig. 4 The ASAP approach to solve BIG DATA Analytics Problems

ISSN: 2456-9348
Vol (02) _Issue (02) Impact Factor: 4.520

iJETRM
International Journal of Engineering Technology Research & Management

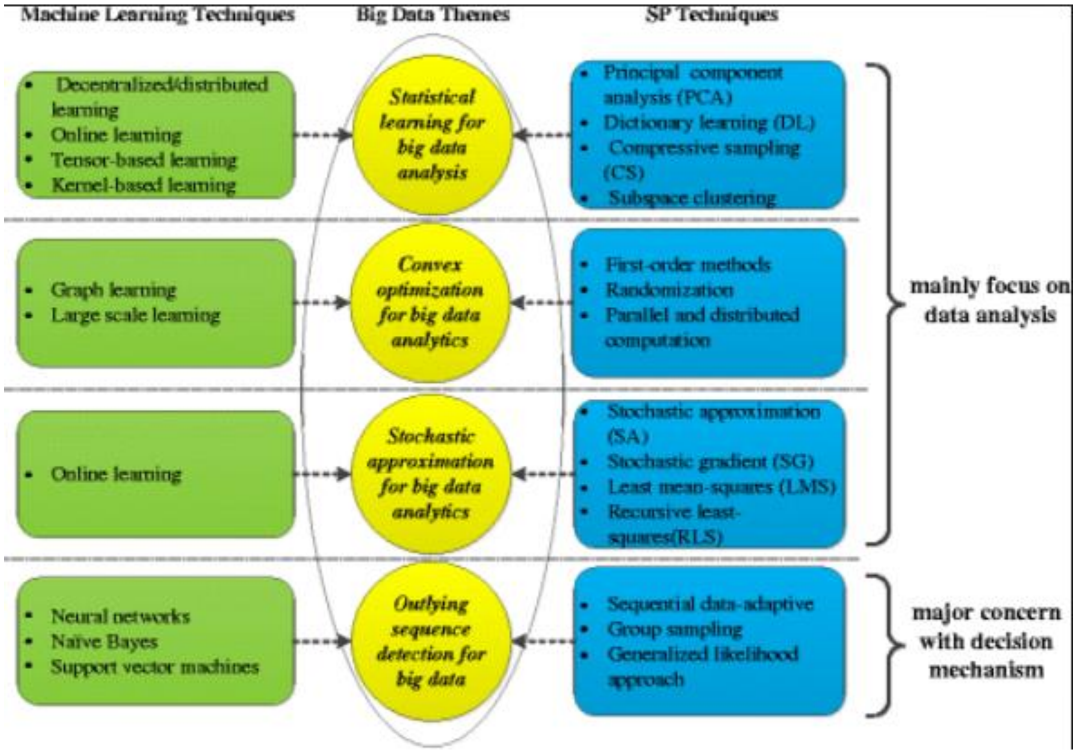


Fig.5 Connection of Deep learning with signal processing techniques for big data from different perspectives

CHALLENGES

Much of this newly available data is in the form of clicks, images, text, or signals of various sorts, which is very different than the structured data that can be cleanly placed in rows and columns. At the same time, we have entered an era when Deep learning can theoretically find patterns in vast amounts of data to enable enterprises to uncover insights that may not have been visible before. Deep learning trains itself on data, and for a time, that data was scarce. Today it is abundant. By 2025, the world will create 180 zettabytes of data per year (up from 4.4 zettabytes in 2013)

Fig.6SummaryofseveralDeeplearningalgorithms

Algorithms	Algorithms	Algorithms	Learningpolicy	Learning	Classification
Decisiontree	Discriminan	Classification	Regularizedmaximum	Feature selection,	IF-THEN rule According to
Non-linear SVM (basedo n	Discriminan	Super-plane separation, kernel trick	Minimizing regular hingeloss softmargi	Sequential minimal optimization	Maximumclasso f testsamples
Linear SVM (basedo n	Discriminan	Super- plane	Minimizingthelossof regular hinge, soft	Sequential dual method	Maximum weighted test sample
Stochastic gradient	Discriminan	Linear combination of weak classifier	Additionminimization loss	Stochastic gradient	Linearcombination of weighted maximum weak classifiers
NaiveBayesian classifier	Generative	Joint distribution of class and feature, conditional	Estimation of maximum likelihood, Maximum	Probabilistic computatio	Maximum posterior

ISSN: 2456-9348
Vol (02) _Issue (02) Impact Factor: 4.520
iJETRM
International Journal of Engineering Technology Research & Management

Big Data and Deep learning would seem to be a perfect match, coming together at just the right time. But it's not that simple.The connected world is ever-widening, enabling the capture and storage of more—and more diverse—data sets than ever before. Nearly 5,000 devices are being connected to the Internet every minute today; within ten years, there will be 80 billion devices collecting and transmitting data in the world. Voice, facial recognition, chemical, biological, and 3D-imaging sensors are rapidly advancing. And the computing muscle that will be required to churn through all this data is more readily available today. There's been a one trillion-fold increase in computing power over the past 60 years.But having vast amounts of data and computing power isn't enough. For Deep learning tools to work, they need to be fed high-quality data, and they must also be guided by highly skilled humans. There's also the misperception that having access to all this new data will necessarily lead to greater insight. These problems are not insurmountable. Tools are being developed to help businesses deal with some of the data management blocking and tackling that stands in the way of advanced analytics. If a company has developed a Deep-learning tool for real estate and finance companies that it says can extract unstructured data in 20 different languages from contracts and other legal documents and transform it into a structured, query-ready format.the critical issues of Deep learning techniques for big data from fivedifferent perspectives, as described in Fig. 3, including learning for large scale of data, learning for differenttypes of data, learning for high speed of streaming data ,learning for uncertain and incomplete data, and learn-ing for extracting valuable information from massive amounts of data.

GeneralchallengesaboutDeeplearningare:(i)designingscalableandflexiblecomputational architecturesforDeeplearning;(ii)theabilitytounderstandcharacteristicsofdatabefore applying Deeplearning algorithmsandtools;and (iii) theability to construct, learnandinfer withincreasingsample size, dimensionality, and categories of labels.Thereare manyscaleDeeplearningalgorithms,butmanyimportantspecificsub-fieldsinlarge-scaleDeep learning,suchas large-scalerecommendersystems,natural languageprocessing, associationrulelearning,ensemblelearning,stillfacethescalabilityproblems.

The basic Map Reduce framework commonly providedby first-generation "Big Data analytics" platformslikeHadooplacksanessentialfeatureforDeeplearning.MapReducedoesnotsupport iteration /recursionorcertainkey features required to efficiently iterate"around"aMap Reduce program.Programmersbuilding Deep learningmodelsonsuchsystemshaveto implement loopinginad-hocwaysoutsidethecoreMapReduceframework.Thislackofsupporthasmotivated the recentdevelopmentofvariousspecializedmethodsor libraries tosupportiterative programming onlargeclusters.Meanwhile, recentMapReduceextensionssuchasHaLoop, Twister, and PrItraimatdirectlyaddressingtheiterationoutage in MapReduce. MajorproblemsthatmaketheDeeplearning (ML)methodsunsuitableforsolving big data classificationproblemsare:(i)AnMLmethod thatis trainedonaparticularlabeleddatasetsmay notbesuitableforanotherdataset–thattheclassificationmaynotberobustoverdifferentdatasets; (ii) an MLmethod isgenerally trained using acertainnumberof class typesandthusa large varietiesofclasstypesfoundinadynamicallygrowingdatasetwillleadtoinaccurateclassification results;and(iii)anMLmethodisdeveloped basedona singlelearningtask,andthereforetheyare notsuitablefortoday'smultiplelearning tasksandknowledgetransfer requirementsofBigdata analytics.

Traditionalalgorithms in ML generallydo notscale tobigdata.The main difficultylieswith their memory constraint.Althoughalgorithmstypically assumethattraining data samplesexistinmain memory,bigdatadoesnotfitintoit.Acommonmethodoflearning from alargedatasetisdata distribution.By replacing batchtraining ontheoriginaltraining datasetwithseparated computationsonthe distributedsubsets,one cantrainanalternative predictionmodelata sacrifice ofaccuracy.Anotherapproachisusing onlinelearning,inwhichmemoryusage doesnotdepend ondatasetsize.Bothonlinelearning anddistributedlearning arenotsufficientforlearning from big data streams. There are two reasons. First isthat thedata sizeis too bigtobe relaxedbyeither onlineordistributed learning.Sequentialonline learning onbig data requires toomuchtimefor trainingonasingleDeep.Ontheotherhand,distributedlearningwithabignumberofDeeps reduces the gained efficiencyperDeep and affects the overallperformance. The secondreason isthatcombiningreal-timetrainingandpredictionhasnotbeenstudied.Bigdataisusedafterbeing storedin(distributed)storage;therefore, thelearningprocessalsotendstoworkinabatchmanner.

Scalingup bigdatato properdimensionalityis achallenge thatcan encounterinDeeplearning algorithms;and there are challengesofdealingwith velocity, volume andmanymore forall types ofDeeplearningalgorithms.Sincebigdataprocessingrequiresdecomposition,parallelism,modularityand/orrecurrence,inflexibleblack-boxtypeDeeplearningmodelsfailedinanoutset.

iJETRM
International Journal of Engineering Technology Research & Management

Applying thedistributeddata-parallelism (DDP)patternsinBigDataBayesianNetwork (BN) learningfaces severalchallenges:(i)effectivelypre-processingbigdatato evaluate its qualityand reducethe sizeifnecessary; (ii)designingaworkflowcapableoftakingGigabytesofbigdatasets andlearning BNswithdecentaccuracy;(iii)providingeasy scalability supporttoBN learning algorithms.
Deep learning challengesinbig dataanalytics lie in:incremental learning fornon-stationary data, high-dimensionaldata, and large-scale model. Because high-leveldata parallel frameworks,like MapReduce donotnaturally orefficiently supportmany importantdata miningandDeeplearningalgorithmsandcanleadtoinefficientlearningsystems,theGraphLab abstractionwas introduced.Itnaturally expresses asynchronous,dynamic,graph-parallel computation while ensuring dataconsistency andachieving ahigh degree of parallel performance in the shared-memory setting

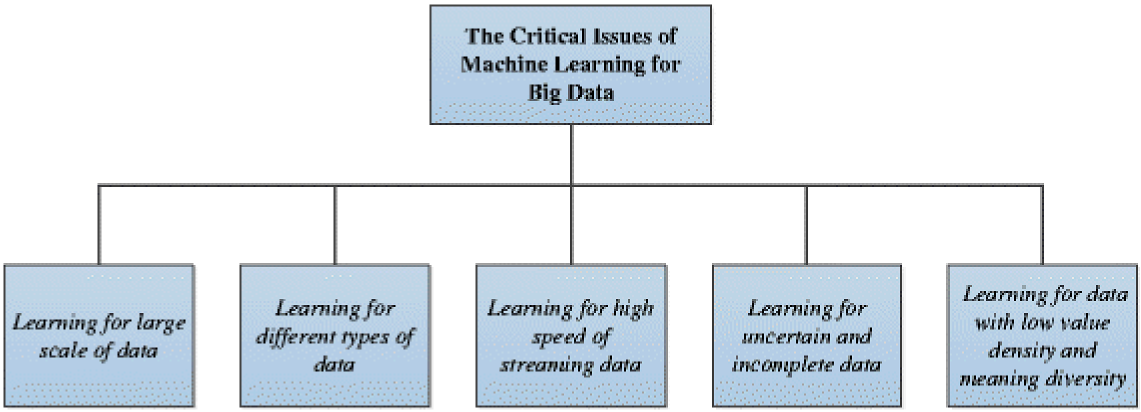


Fig.7 The vital issues of Deep learning for big data

CONCLUSION

The learning comes from extensive calculations done over existing datasets to create a learning model. A normal system can't handle very large dataset calculation and data size is increasing day by day, thus the obtained model should be adapted accordingly. To obtain this we have to implement distributed computing using big data technologies like Apache Mahout, Spark, R-Hadoop or initial analytics processing in projects like hive/ pig and feed output to Deep learning algorithms for model/ learning generation.
Deeplearningapplicationsinbigdatahasmetchallengessuchasmemoryconstraint,nosupport (initerations)fromMapReduce,difficultyindealingwithbigdataduetoVs(suchashighvelocity, volume,andvariety,etc.), and learning training limited to a certain number of class types or a particular labeled datasets,etc .Some technology progress has been made such as faceted learning for hierarchical data structure, multi-task learning ininparallel,multi-domain/cross-domainrepresentation-learning,streaming dataprocessing,high-dimensionaldataprocessing,andonlinefeatureselection, etc. These areas and the above challenges about Deep learning in bigdata also can be further research topics.

REFERENCE

[1] R. Krikorian, Twitter by the Numbers Twitter, 2010, [online] Available: http://www.slideshare.net/raffikrikorian/twitter-by-the-numbers?ref=http://techcrunch.com/2010/09/17/twitter-seeing-6-billion-api-calls-per-day-70k-per-second/.
[2] Billion Devices Will Wirelessly Connect to the Internet of Everything in 2020, 2013, [online] Available: https://www.abiresearch.com/press/more-than-30-billion-devices-will-wirelessly-conne/.
[3] W. Raghupathi,W&Raghupathi, V (2014) Big data analytics in healthcare: Promise and potential, Health Inf. Sci. Syst., vol. 2, pp. 1-10.
[4] Bach. F. (2011) Learning with Submodular Functions: A Convex Optimization Perspective. URL http://hal.inria.fr/hal-00645271/en
[5] Bach,F ,Jenatton, R., Mairal, J. and G. Obozinski (2011) Optimization with sparsity-inducing penalties.Foundations and Trends R in DeepLearning, 4(1):1–106.

iJETRM
International Journal of Engineering Technology Research & Management

[6] Blatt, D., Hero, A.O. and Gauchman, H.(2008). A convergent incremental gradient method with a constantstep size. 18(1):29–51.
[7] Jenatton, R, Obozinski, G and Bach, F(2009) Structured sparse principal component analysis. Technicalreport, arXiv: 0909.1440.
[8] Jenatton, R., Mairal, J., Obozinski, G and Bach, F(2011) Proximal methods for hierarchical sparse coding. Journal of Deep Learning Research, 12:2297–2334.
[9] N. Le Roux, M. Schmidt, and F. Bach. A stochastic gradient met hod with an exponential convergence rate for strongly-convex optimization with finite training sets. Technical Report -, HAL, 2012.
[10] Mairal, J, Bach, F. Ponce, J and G. Sapiro G (2009). Online dictionary learning for sparse coding. In International Conference on Deep Learning (ICML) , 2009.
[11] ObozinskiG. and Bach.F (2012) Convex relaxation of combinatorial l penalties. Technical report, HAL,
[12] JunfeiQiu, Qihui Wu, Guoru Ding, Yuhua Xu and Shuo Feng (2016)A survey of Deep learning for big data processing EURASIP Journal on Advances in Signal Processing 2016:67 DOI 10.1186/s13634-016-0355-x