iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 5 **Issue:** VI **Month of publication:** June 2017

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

Volume 5 Issue VI, June 2017
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Understanding the Fundamentals of Machine Learning Field, Its Evolution & Its Need

Manish Kumar Singh[1], Prof. G S Baluja[2], Dr. Dinesh Prasad Sahu[3]
Baluja Institute of Technology& Management
DTTE, The Government of NCT & Delhi
Jawaharlal Nehru University

Abstract: ***Since the time of inception of computers, many computer scientists and researchers have been fascinated by the idea of programming machines in such a manner that can make them learn the way humans do from their surrounding and past experiences without any explicit source to control them. This envisage had led to the beginning of machine learning field and the artificial intelligence played a crucial role in the inception of machine learning. With the advancements made in the field of machine learning concepts and algorithms such as neural network, decision tree, deep learning, cloud computing and Big Data analytics, the entire humankind has experienced huge developments in various industries and research field. The progress experienced in the twenty-first century in all kinds of human endeavors including - web search, marketing & sales, health care, bioinformatics, space research, oil & resource exploration transportation and much more, machine learning has undeniable contributions to such progress. This paper contains the relevant review of machine learning fundamentals, its evolution and its need in the present era as well as in the upcoming future.***
Keywords: ***Machine Learning (ML), Supervised Learning, Unsupervised Learning, Reinforcement Learning***

I. INTRODUCTION

Today's information age is unarguably the period with most defining discoveries and developments in the history of the mankind. And the key development area that has regulated this period is the Internet. It has seen vast development since the time of its inception. As the time passed, the cyber world swayed from one invention to the next with each invention bringing the new areas of research in the front of computer scientists and cyber enthusiasts. Out of these areas of research, Machine Learning (ML) has proved to be one of the most envisaged fields of research for quite a long period. And the key idea for doing research in Machine Learning is to provide the computers enough capability to learn somewhat how humans do so that they do not need any human assistance to work with them. The idea about developing self-automated machines is the key concern of the Machine Learning field that enables machines to learn by themselves automatically so that their knowledge about their surrounding can be improved by themselves through each passage of time. Such way of learning capability acquired by the machines can definitely help them to recognize and understand the data to make a good decision by them on the basis of the supplied data. Machine Learning involves the study and development of myriad algorithms to help machines to make different kinds of decisions on the basis of different supply of data. Such algorithms have the enough impact on machines to develop knowledge out of the provided data by making the use of possible statistics, logic, probability theory, search, control theory, combinatorial optimization and reinforcement learning.

This article is an effort to provide key understanding to the machine learning field, its need and its evolution. The section 2 of this article analyses the fundamentals of machine learning concepts while The section 3 of this article gives an anecdote about the evolution of machine learning field. The section 4 put deep emphasize on the need of the machine learning knowledge with respect to various areas of research. and the section 5 wraps up the article with a suitable conclusion.

II. UNDERSTANDING THE FUNDAMENTALS OF MACHINE LEARNING FIELD

Machine learning provides a marvelous way to do data analytics that helps to automate the building of the analytical model. It provides various algorithms that help machines to learn from data so much so that they can search for hidden insights without letting them be programmed explicitly to do so.

Machine learning helps the machine to solve the real-world problems that generally of highly complex nature. The problems of any real–world kind such as – "Is he suffering from cancer?", "What is the cost of this fruit?", "Who are your nice friends?", "Will the given deal be profitable?", "What did you tell me last Sunday?", "What is my cell phone number?" etc. may be solved dramatically through the algorithms defined under machine learning. But such algorithms can be worked upon only after having a deep

©IJRASET: All Rights are Reserved

Volume 5 Issue VI, June 2017
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

understanding for the different types of machine learning concepts that can be distinguished from one another in the following manner:

A. Supervised learning

Under this machine learning type, the machine learning algorithms provide learning training through labeled examples, somewhat like that of an input for which the output is known beforehand. For instance, the input like this "which is the largest continent of the world" has the output already known as "Asia". The algorithms working under supervised learning takes the inputs for which the output is already known for the reason so that the algorithms can make the machine to learn by letting it compare the actual output with the already known output to check for any further errors. The machine is then modeled accordingly.

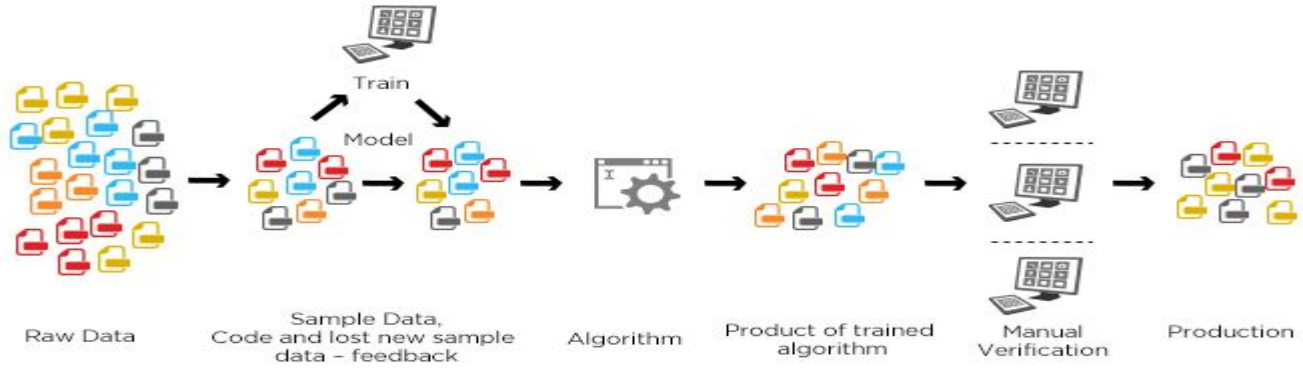


Fig. 1 Supervised Learning

The famous supervised learning algorithms include classification, gradient boosting, prediction and regression. It then modifies the model accordingly. With such algorithms, a machine makea use of supervised learning to do the prediction of label values on unlabeled data by using suitable patterns. Supervised learning finds the application in such areas where the future events are predicted through the historical data. For example, a machine can help to predict through supervised learning that which of transactions made through a credit card can be fraudulent. Some common examples of supervised learning include classification of e-mails for checking them for being spam or not, voice recognition and labeling of web pages on the basis of their content.

B. Unsupervised learning

Under this machine learning type, the machine learning algorithms are made to work with data which may not have any historical labels. In order to do so, the algorithm must be worked upon to find out what is being actually presented to learn for the purpose to look out data so that some structure can be determined within. Such kind of learning type is quite useful while working with transactional data. Some common examples of unsupervised learning include the identification of a group of customers on the basis of similar attributes that can be targeted during digital marketing campaigns done by the online retailers.

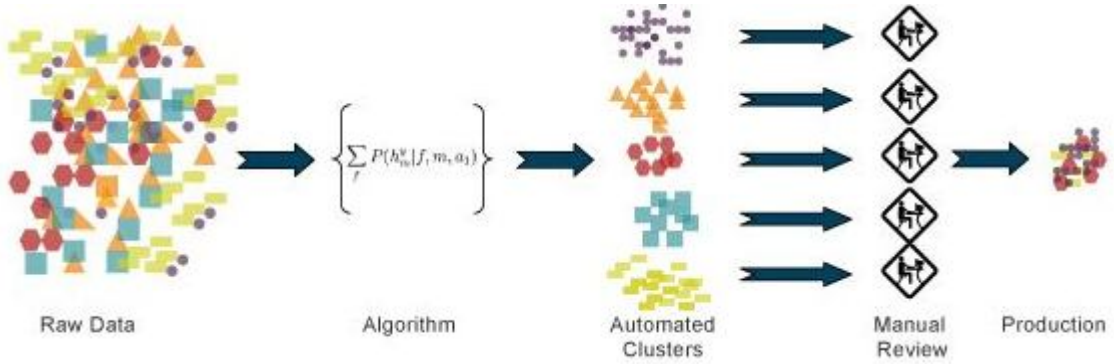


Fig. 2 Unsupervised Learning

The famous unsupervised learning algorithms are k-means clustering, self-organizing maps, hierarchical clustering, nearest neighbor mapping etc. The unsupervised learning algorithms are quite helpful to group text on the basis of given topic, recommend online

©IJRASET: All Rights are Reserved

Volume 5 Issue VI, June 2017
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

available items and the identification of data outliers.

C. Semi-supervised learning

Under this machine learning type, the machine is made capable of learning both labeled and unlabeled data for the training purpose. This especially involves training the machine through a small amount of labeled data at the expense of a large amount of unlabelled data. This is for the reason that unlabelled data are economical and easy to gather. This type of machine learning is used frequently with the algorithms like classification, prediction and regression. Further, this type of learning is used in the field where the cost of an associated labeling is splurging to make way to a fully labeled training process. The famous application of semi-supervised learning is face recognition through a webcam.

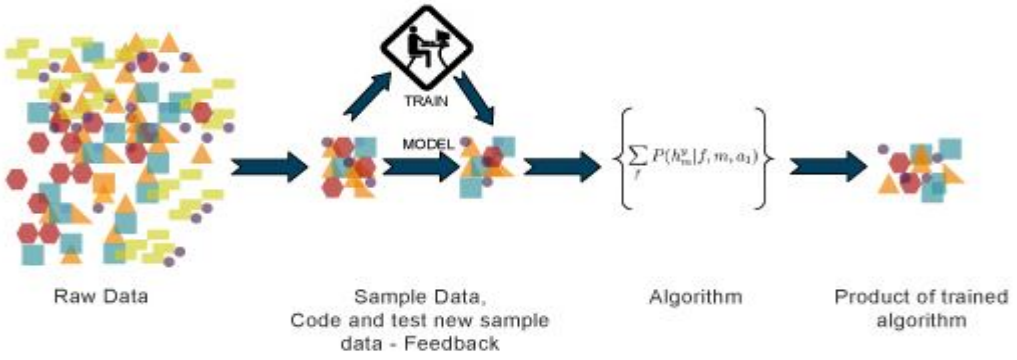


Fig. 3 Semi-supervised Learning

D. Reinforcement learning

Under this machine learning type**,** the machine learning algorithms work through the trial and error approach to make sure of the actions that give the best results and it founds applications in the field of gaming, navigation, and robotics. is often used for robotics, gaming, and navigation. There are three components that work primarily under this machine learning type - the agent/decision-maker/learner, the environment with which the agent do the interaction and the actions that the agent is supposed to do.

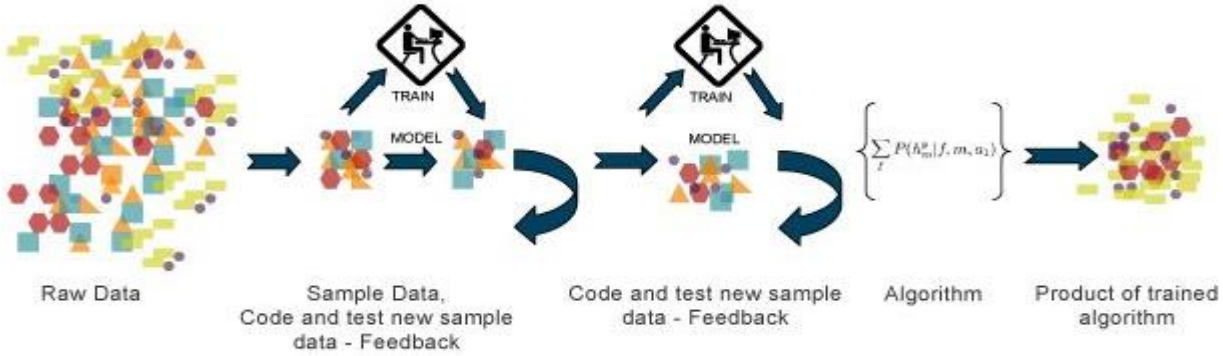


Fig. 4 Reinforcement Learning

The entire objective of reinforcement learning is to make the agent select actions that can help to obtain maximized reward over the specified period of time. So the idea is clear that the reinforcement helps the machine learn the best policy to work with to give best results.

III.MACHINE LEARNING EVOLUTION

The machine learning field came into existence after following the quest for artificial intelligence field. It has been analyzed through the media prints and historic archives that the desire of developing a machine that could be intellectually capable as much as humans had fired the imagination of writers and also early computer scientists to give rise to the machine learning field.

A. The first machine learning based system was developed by Arthur Samuel in 1952. At IBM, he developed a program for playing checkers. The program was able to observe the positions at the game and learn a model that gives better moves for the machine player. The system played many games with the program and observed that the program was able to play better in the

©IJRASET: All Rights are Reserved

course of time with getting more experience of more games. Samuel coined the term "Machine Learning" and he defined the learning as "a field of study that gives a computer the ability without being explicitly programmed".

B. In 1957, Rosenblatt proposed the perceptron – a simple neural unit. It was a very exciting discovery at that time. Rosenblatt made the statement about perceptron that the perceptron was designed to illustrate some of the fundamental properties of intelligent systems in general, without becoming too deeply involved in the special and the frequently unknown conditions which hold for particular biological organisms. But after three years, Windrow and Hoff came up with the Delta Learning Rule that is used for learning perceptron. It was used as a procedure for training perceptron. It is also known as the least mean square (LMS) rule. The combination of theses ideas created a good linear classifier. However, the work along this line suffered an as big setback when Minsky came up with the limitations of perceptron in 1969. He showed that the XOR problem could not be represented by perceptron and such inseparable data distribution cannot be handled.

C. Following Minsky's work, neural network research went dormant up until the 1980s. In the meantime in the 1970s, machine learning was done on the basis of the algorithms related to Symbolic artificial intelligence and Good Ol' Fashioned artificial intelligence where concept inductions were worked on to do machine learning. John Ross Quinlan came up with Decision Tree learning – specifically the ID3 algorithm in 1986. ID3 was also released as a software which had simplistic rules in contrary to the black box of the neural network which became quite popular. After ID3, many alternatives or improved ID3 were developed such as CART and Regression Trees. They are still one of the most popular topics of machine learning. During this time, Symbolic Natural language processing also became very popular.

D. In the 1980s, Advanced decision tree and rule learning were developed. Learning, planning, and problem-solving were there and at the same time, a resurgence of the neural network was there. The intuition of multi-layer perceptron was suggested in 1981 and neural network specific Back-Propagation algorithm was developed. Back-Propagation is the key ingredient of today's neural network architectures. With those ideas, neural network research became popular again and there was an acceleration in 1985-86 when neural network researchers presented the ideas of MLP i.e. Multi-Layer Perceptron with the practical PP training. David Rumelhart,Hanton Williams, and H Wilson were the important scientists who worked n this area. During this time, a theoretical framework of machine learning was also presented. For instance, Valiant's PAC learning theory which stands for probably approximately correct learning was developed and the focus was then shifted to experimental methodologies.

E. In the 1990s, machine learning embraced statistics to a larger extent. It was during this time that the Support Vector Machines were proposed. It was a machine learning breakthrough and the Support Vector Machines was proposed by Vapnik and Corinna Cortes in 1995. Another strong machine learning algorithm was proposed by Yoav Freund and Robert Schapire in the form of *AdaBoost*, short for "Adaptive Boosting by which they could create strong classifiers from ensembles of weak classifiers. The Kernelized version of SVM dual was proposed in the 2000s and it was able to exploit the knowledge of convex optimization, generalization, and kernels. Another ensemble model was explored by Bryan in 2001 that ensembles multiple decision trees where each of them is curated by a random subset of instances. This is called a random forest. During this time, Bayes Net learning was also proposed. Then neural network research suffered another damage by the work of Hawkwriter who showed that the gradient loss after the saturation of neural network unit happens when we apply back-propagation so that after certain numbers of epochs the neural networks are inclined to over fit.

F. But, as we come closer today, we see that the neural networks are again very much popular. We have a new era in a neural network called deep learning and this phrase refers to a neural network with many deep layers. This rise in neural networks began roughly in 2005 with the conjunction of many different discoveries of people like Hilton, Lejeune, Bangor, Andrew MG and other researchers. At the same time, let us look at certain applications where machine learning has come to the public forefront:

1) In 1994, the first self-driving car made the road test;
2) In1997. Deep Blue beat the Grandmaster Gary Kasparov in the game of chess;
3) In 2009, we saw Google building the self-driving car;
4) In 2011, Watson, again from IBM, won the popular game of Jeopardy;
5) In 2014, we saw Human Vision surpassed by the ML systems; and
6) In 2014-15, we found that the machine translation systems driven by neural networks proved to be better than other statistical machine translation systems;\

©IJRASET: All Rights are Reserved

G. There are certain concepts and certain technologies that are making headlines in machine learning from time to time which include GPU's (that enables the use of machine learning and deep neural network), the Cloud technology, the availability of Big Data – all these are making the concept of machine learning quite interesting.

IV. THE NEED OF MACHINE LEARNING KNOWLEDGE

Machine learning finds major application in those areas which have to deal with huge amounts of data on daily basis. The industries belonging to such areas work on such data one way or another through machine learning algorithms to derive important information out of them that help in good decision making, profit earning and overriding the competitors in the market. The key roles played by machine learning in some of the important areas are as follows:

A. Providing Services to Financial Institutions

The financial institutions like bank, insurance companies, foreign exchange etc. make use of machine learning based technologies and tools for two of the key purposes – (a) to derive relevant information from Big Data for good decision making by going into deep insights of the data, and (b) to counter the fraud like activities. The former purpose of going into deep insights of the Big Data can be done through the identification of opportunities for doing investments, or providing help to the investors to know when and how to do trade. The process like data mining play a crucial role in this regard that makes a way to do identification of such clients who bear high-risk profiles or it can help to figure out the warning signs of fraud through cyber-surveillance.

B. Aiding Government Agencies

Machine learning provides great aid to government agencies to ensure public safety and utilities by helping the agencies to derive important information from mined data after going deep insights of them. For example: machine learning algorithms can be used to do an analysis of sensor data to identify the ways that can increase security based services at the expense of less money spending. Further, detection of fraud and minimization of the identity theft can easily be done through machine learning. Image recognition and pattern recognition are used in Bioinformatics to ensure security and tracking down the activities of the concerned person in an organization.

C. Contributing a Huge to Healthcare Industry

Healthcare industry has gained at large from machine learning based tools and technologies. For instance, the discovery of wearable sensors and devices through machine learning algorithms have helped the doctors and path labs to do real time checking of a patient after going into the insights of data collected from the tests done on the patient. Moreover, the identification of trends or ways to diagnose disease and do its treatment has got a major boost through machine learning. For example, MYCIN is an expert system developed through machine learning that helps to do identification of bacteria, such as bacteremia and meningitis, that causes severe infections, and it helps to recommend corresponding antibiotics with the knowledge of adjusted dosage as per body weight of a patient.

D. Doing Digital Marketing to Improve Sales

Machine learning based tools in build in the websites provide recommendation of items to customers on the Internet that they may like. Such recommendations of products are made by websites by tracking down the previous purchases of the customers and do analysis of their buying history through machine learning algorithms. After obtaining the relevant information through analysis, companies and online retailers do web personalization and digital advertising to promote their products on the Internet.

E. Exploration of oil and Natural Gas

Machine learning based tools and technologies have worked wonderfully for the years to explore the areas and regions with the abundance of oil and natural gas resources. The main tasks of machine learning algorithms under the exploration of oil and natural gas include the search for latest energy sources, prediction of the failure of refinery sensor, analysis of minerals found in the ground, streamlining for the distribution of oil to make it more economical and efficient to access.

F. Contributing the Transportation Industry

The transportation industry makes use of machine learning to do identification of patterns out of Big Data to make way to derive the best possible information related to efficient route coverage that could make the transportation profitable. This way of working with

©IJRASET: All Rights are Reserved

machine learning tools and technology has contributed hugely to the transportation of raw materials, carriage products, people etc. from one part of the world to another.

V. CONCLUSIONS

Machine learning is undoubtedly the leading research field in computer science field today. It has contributed a huge to the progress and development made in various spheres of mankind. The concepts of machine learning like supervised learning, unsupervised learning and reinforcement learning are very necessary to build the tools and technology that can work effectively to derive relevant information from the deep insights of Big Data. However, the journey of machine learning field has not been so easy and straightforward. It took a lot of pain and trouble of various researchers & computer scientists and it took around 70 years to make the field of machine learning so exciting and fruitful as it looks today. No area of human endeavor has left where machine learning concepts are not contributing. But the research scholars and the computer scientists need to focus more on the areas of deep learning, cloud computing and Big Data Analytics to obtain further and deeply hidden benefits and scope of machine learning.

VI. ACKNOWLEDGMENT

We would like to acknowledge all the authors, researchers and the data scientists whose works that have been mentioned in the references had helped us a huge to come up with this paper to analyse the machine learning field. We would also like to say thanks to the journal – IJRASET and its most respectable Editor-in-chief which has given a space to our paper in their esteem journal. .

REFERENCES

[1] Prof. Sarkar, Sudeshna. Introduction to Machine Learning, NPTEL. 4 June 2016. Web. https://www.youtube.com/watch?v=T3PsRW6wZSY&list=PLlGkyYYWOSOsGU-XARWdIFsRAJQkyBrVj
[2] Machine Learning: What it is and why it matters, SAS. 4 June 2016 Web. https://www.sas.com/en_us/insights/analytics/machine-learning.html
[3] Ray, Sunil. Essentials of Machine Learning Algorithms (with Python and R Codes), 10 Auhgust 2015. Web. https://www.analyticsvidhya.com/
[4] Machine learning. Wikipedia. Web. https://en.wikipedia.org/wiki/Machine_learning
[5] Mahout - Machine Learning. Tutorial Point. Web. https://www.tutorialspoint.com/mahout/mahout_machine_learning.htm
[6] Mccrea, Nick. An Introduction to Machine Learning Theory and Its Applications: A Visual Tutorial with Examples. Toptal. Web. https://www.toptal.com/machine-learning/machine-learning-theory-an-introductory-primer
[7] Ron Kohavi; Foster Provost (1998). "Glossary of term"". Machine Learning. **30**: 271–274.
[8] Dickson, Ben. "Exploiting machine learning in cybersecurity". TechCrunch. Retrieved 2017-05-23.
[9] Wernick, Yang, Brankov, Yourganov and Strother, Machine Learning in Medical Imaging, IEEE Signal Processing Magazine, vol. 27, no. 4, July 2010, pp. 25–38
[10] Mannila, Heikki (1996). Data mining: machine learning, statistics, and databases. Int'l Conf. Scientific and Statistical Database Management. IEEE Computer Society.
[11] Friedman, Jerome H. (1998). "Data Mining and Statistics: What's the connection?". Computing Science and Statistics. **29** (1): 3–9.
[12] "Gartner's 2016 Hype Cycle for Emerging Technologies Identifies Three Key Trends That Organizations Must Track to Gain Competitive Advantage". Retrieved 2017-04-10.
[13] "Why Machine Learning Models Often Fail to Learn: QuickTake Q&A". Bloomberg.com. 2016-11-10. Retrieved 2017-04-10.
[14] Mitchell, T. (1997). Machine Learning. McGraw Hill. p. 2. ISBN 0-07-042807-7.
[15] Simonite, Tom. "Microsoft says its racist chatbot illustrates how AI isn't adaptable enough to help most businesses". MIT Technology Review. Retrieved 2017-04-10.
[16] Harnad, Stevan (2008), "The Annotation Game: On Turing (1950) on Computing, Machinery, and Intelligence", in Epstein, Robert; Peters, Grace, The Turing Test Sourcebook: Philosophical and Methodological Issues in the Quest for the Thinking Computer, Kluwer
[17] Fig-1, Explore Supervised Learning, Machine Learning, and more!, Simplilearn. Web. https://www.pinterest.com/pin/104427285088231073/
[18] Fig-2, Explore Supervised Learning, Machine Learning, and more!, Simplilearn. Web. https://www.pinterest.com/pin/104427285088231073/
[19] Fig-3, Explore Supervised Learning, Machine Learning, and more!, Simplilearn. Web. https://www.pinterest.com/pin/104427285088231073/
[20] Fig-4, Explore Supervised Learning, Machine Learning, and more!, Simplilearn. Web. https://www.pinterest.com/pin/104427285088231073/
[21] Fig-5, Explore Supervised Learning, Machine Learning, and more!, Simplilearn. Web. https://www.pinterest.com/pin/104427285088231073/

ABOUT AUTHORS

Manish Kumar has published two research papers in international journals on web mining – chiefly "Web Mining: Penning an Era of Information Age" and "Understanding How Crucial Hidden Value Discovery In Data Warehouse Is?" and he has also published three articles in international journal. His areas of interest include Big Data Analytics and Machine Learning.

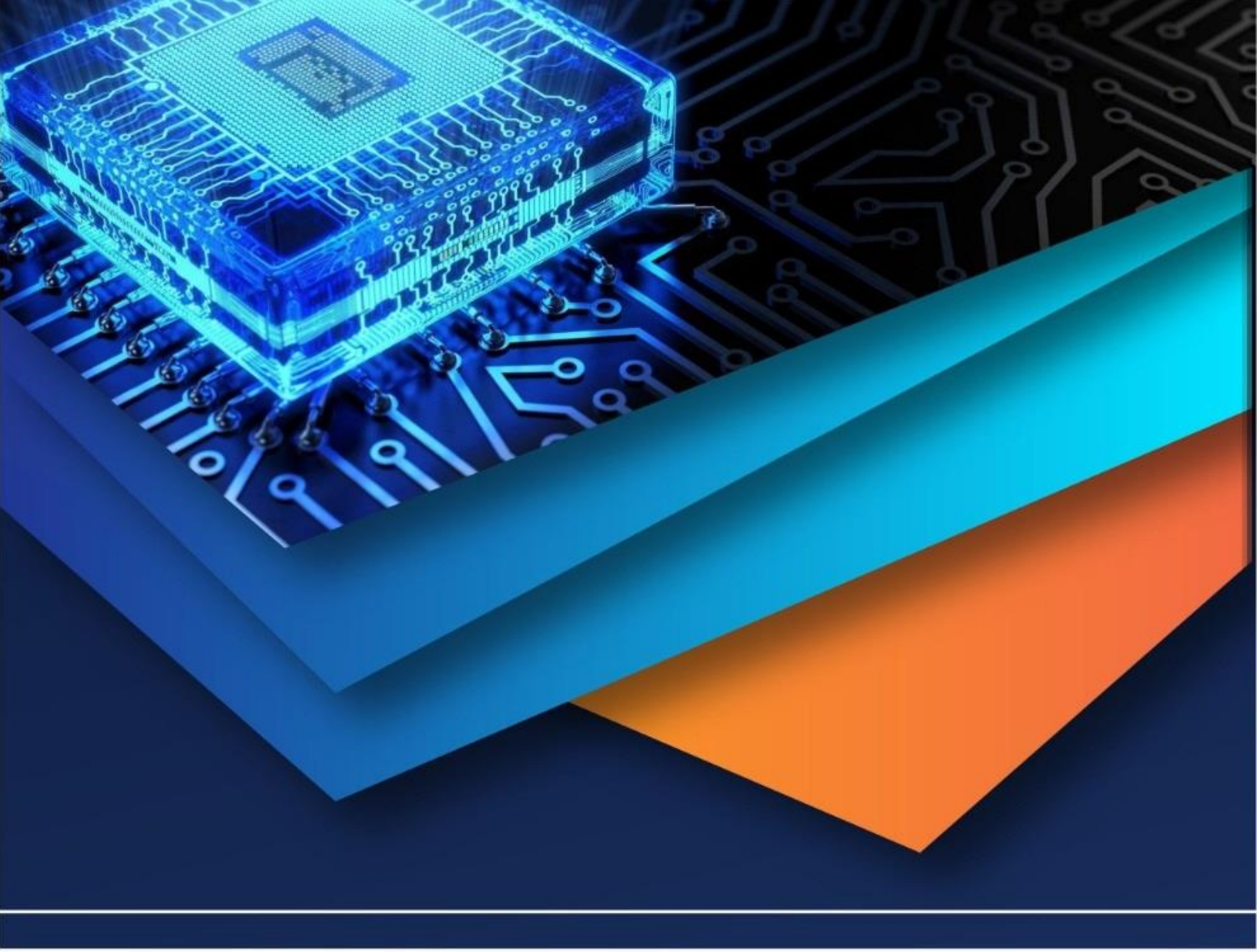
©IJRASET: All Rights are Reserved

Volume 5 Issue VI, June 2017
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

GS Baluja is a famous Indian author of Computer Science field who has authored numerous books so far - chiefly Data Structure Through C, Data Structures Through C++, Object Oriented Programming Using C++, Java Programming & many more. He has done B.E (Com. Sci.) from Marathwada University and M.Tech from IIT, Delhi.

Dinesh Prasad Sahu received the Master degree (Computer Science & Application) M.Tech (Computer Science & Application) from Jawaharlal Nehru University, New Delhi, India. Currently, He is doing Ph.D. (Computer Science & Engineering) under the guidance of Dr. Karan Singh, from Jawaharlal Nehru University, New Delhi, India & is working in school of Computer & Systems Sciences, Jawaharlal Nehru University, New Delhi. His primary research interests include parallel and distributed system and Grid Computing. He has published 3 papers in proceedings of peer-reviewed Conferences.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)