52 pages including cover

CSI Communications

Knowledge Digest for IT Community

Volume No. 42 | Issue No. 8 | November 2018

₹ 50/-

ISSN 0970-647X

www.csi-india.org


Computer Society of India™
Wishing you a
Diwali
that lights up with
Joy • Happiness • Peace & Prosperity

CSI COMMUNICATIONS

VOLUME NO. 42 • ISSUE NO. 7 • NOVEMBER 2018

Contents

Printed and Published by Akshaya Kumar Nayak on behalf of Computer Society of India, Printed GP Offset Pvt. Ltd. 269, 2nd Floor, A-2, Shah & Nahar Indl. Estate, Sitaram Jadhav Marg, Lower Parel, Mumbai 400 013 and Published from Samruddhi Venture Park, Unit No. 3, 4th Floor, Marol Industrial Area, Andheri (East), Mumbai 400 093. • Email : **hq@csi-india.org**
Editor: Prashant R. Nair

Editorial

Prof. (Dr.) S. S. Agrawal
Chief Editor

Dear Fellow CSI Members,

"McKinsey predicted that by 2018 there will be a shortage of 1500,000 data experts "Average salary of Big data Hadoop developers is USD 135000".

It is common challenge for decision makers: how do we make optimal choices with so many unknown variables? It turns out that business insights come from turning what is unknown into what is known. CIO's need to think about information as the new oil 'Information has unique economic characteristics that render it potentially much more valuable to their business than any other fossil fuel.

The current issue of CSI Communications is on BIG DATA- An area of emerging importance and interdisciplinary area of research and applications. The issue is enriched by a number of articles received from professionals, researchers as well as business communities.

Big Data combined by topics such as data analytics, and machine intelligence is a technology-driven process for analyzing data and presenting actionable information to help executives, managers and other corporate end users make informed business decisions. A lot of work has to be done and applied for numerous applications.

The cover story BIG DATA – The Road Ahead is written by Prof. Sunil Gupta and Prof. Goldie Gabrani – well known experts in this area from BML Munjal University . They have high lighted the important characteristics, tools & technologies and its various applications in business. Currently organizations are starting to see that data and content should not be considered separate aspects of information management, but instead should be managed in an integrated enterprise approach. Enterprise information management brings Business Intelligence and Enterprise Content Management together.

Current Trends and Research Front:

On the research front an article on Persistent Speech Recognition Using Deep Learning and several articles technical nature make this issue very important. Some of these include Big Data in Business intelligence, and in Agricultural field, Big Data Computing, Big Data Analytics, etc.

Articles of general nature include An Analysis of Data Mining using Python, Use of ICT in Education.

The issue include notice about Call for nominations by the CSI National Nomination Committee for Elections 2019-2020/2021. The last date for submission of nominations is November 17, 2018.

This issue also gives detailed information about the ICACCP-2019 Project Contest and INDIACom-2019. We encourage you to prepare for the same and actively participate in it. Information about the activities that have taken place at various regions and divisions and the students chapters is also given.

Very nice and informative reports about regional, chapters and student chapter activities. Congratulations on various achievements.

We are thankful to all the contributors and look forward to receive your valuable articles in future also. We express our gratitude to all the ExecCom members and the CSI Officials. We look forward to receive constructive feedback and suggestions from our esteemed members and reader of CSI fraternity. Please log on to http://www.csi-india.org/ and email to csic@csi-india.org.

With kind regards,

Prof. (Dr.) S. S. Agrawal
Chief Editor

Message from the Vice President cum President Elect

From : Vice President, Computer Society of India
Date : 01 November, 2018
Email : vp@csi-india.org / Cell : (91) 82106 93239

The theme of this issue of CSI Communication **Big Data Analytics** is of great importance as it will focus on Technology Innovation and Trend Setting initiatives in Academic, Research, Corporate, Business, Industries, Government, Education, Security and Health Care domains for the citizens. The Global economy will experience the contribution of this great Technology in the current decade but the benefits of the same should be completely and uniformly understood and utilized by the Professional Community.

In recent years, CSI has played a significant role in creating IT awareness by organizing massive numbers of Seminars, Conferences & Workshops starting from Chapter Level, State Level, National Level & International Level on current state of affairs of technological development. The year 2017-18 has witnessed more than 400 events which was mentioned in last year Annual Report. This year the quantum of such activities are in progress & such events are regularly covered in CSI communication.

I request to the Regional Vice Presidents, Chapter Office Bearers and Student Branch co-ordinators to kindly organize more & more activities to take forward the name & fame of CSI to common masses with their realization of benefits from this largest Society of IT professionals.

The birth of CSI was inspired by the pioneers of Information Technology in India. Started with 16 members in 1965. CSI has scaled it's membership height to more than lakhs. It is not only the reflection of the increasing outreach of CSI but also expanding the boundaries to the remotest part of the country with the faith of the members in the Society. The strength of CSI lies in its appeal & association with Students, Academicians, Scientists, Computer and IT Professionals as well as practitioners.

Though CSI is having good numbers of Chapters, Institutional Members, Students Branches still then it's association with the increasing number of companies & corporations are most required. In this direction, the efforts will be to enhance our services so as to render satisfactory services to the members & ensure their continued association.

In spite of the problems, difficulties & misguiding attempt of some of the law deviating persons, the membership of CSI is continuously increasing, particularly the high growth of student membership as reported by the Education Directorate recently. Recently The Chapter shares for 3rd & 4th quarter of 2016-17 (which were long due) have been released to the Chapters, so that they can be able to conduct more activities. The calculation of Chapter shares for the year 2017-18 is under process & the same shall be released just immediately once the process is complete.

Every member of CSI knows that csi-india.org is the official website of CSI since the development of website by the Computer Society of India. A couple of Persons are misguiding the members through an unauthorized website (for which case has already been registered in Cyber Crime Police Station) & giving the wrong information. Law will take it's own action. Therefore if somebody is having the faith in law, there is no need of interpreting his own as correct & mislead the members.

CSI is functioning well. Technical activities are conducted by the chapters & student branches which are quite visible by the reports published in CSI Communication. Chapters & student branches are requested to come with more & more technical activities to cater the effective services to their members, because it is the need that members should be addressed regularly & value to members should be enhanced.

I seek the active & kind support of the Members to make CSI more Dynamic, Vibrant, Productive & Sustainable to achieve the height of excellence.

I sincerely request all the Office Bearers, Executive Members, Chapter Managing Committee, CSI student Branch Coordinator, SIG Managing Committee & CSI Office Staffs to kindly work with responsibility for the Society (CSI) to serve honestly for the cause of every Division, Region, Chapter, SIG, Student Branch & every Individual Members including Student Members.

Let us come forward to make Clean CSI & Green CSI with transparent activities & visions to make it Swachh, Pardarshi & Hara Vara.

Wishing all of you a Happy & Prosperous Deepawali.

AkNayak

Prof. Akshaya Nayak
Vice President, CSI

Big Data - The Road Ahead

Sunil Gupta and Goldie Gabrani
Department of Computer Science and Engineering, BML Munjal University, Gurgaon, India

Introduction

Recent advances in digital technologies have escalated the splurge of data. A large volume of data is being generated in today's world in form of semi-structured, structured and unstructured form. These data sets are referred to as big data. Today, almost every domain in the world is data driven. In the field of science, there are data sets related to automobiles, pollution, genomics, environmental, astronomy etc. Social Sciences and humanities have historical documents, scanned books and social interactions data. Corporate has data related to sales, supply chain, stock market transactions, rail traffic, banking transactions. Entertainment has music, videos, internet images, Hollywood movies, and world of medicine has MRI & CT scans, patient health records, blood bank, universities have variety and large amounts of research data and technical literature. In addition lots of data is being generated on regular basis in big magnitudes, every instant of time in variety of forms (numbers/text/audio/images/video) from multiple sources including systems, sensors, mobile devices, digital process and social media. As examples, Google processes approximately 20 Peta byte of data in a day, wayback system has three Peta byte plus 100 Tera Byte (TB) per month, Facebook has 2.5 Peta byte of user-data plus 15 Tera byte data per day, 12 Terabytes of Tweets created each day, eBay has 6.5 Peta Byte of user data plus 50 Tera byte per day and CERN's produces 15 Peta Byte data in a year [1].

The key enablers that have given this great rise in the growth of big data have increased along with availability of cheap storage capacity, enhanced processing power and methods to access the data. As the data that is being generated has an alarming velocity, volume and variety; the data sets thus produced are so huge and it becomes a challenging task to process them using conventional processing tools and techniques of data.

This challenge was first faced by organizations that were in the area of developing web search algorithms. These organizations were faced with the issues of querying very large accumulations of loosely structured data that too distributed in nature. Google then came up with an innovative solution to this by developing Map Reduce. Map Reduce supported distributed computing environments on large data sets residing on cluster of computers. Since then, Big Data is being used by large number of organizations to improve the quality and value of their products and services [1].

Big Data and some related terms

Data exists in variety of forms including relational data, text data, structured and semi-structured data, graphical data, semantic web, social network, and streaming data. It consists of facts and statistics collected together in order to do analysis or for simply for reference. One can do aggregation and Statistics like data warehousing, searching, indexing, and querying based on keyword and pattern matching and knowledge discovery of data.

Big data is huge amount of unstructured, semi structured and structured data; which grows at very high rate with regular or irregular peaks and dips, having variety of forms, and sometimes hard to associate with other data sets. Big data does not specify amount of data that defines Big Data. In fact it represents terabytes and petabytes of data that has been collected over a period of time. Big data is the technology that faces challenges in terms of Volume, Communication, Speed of Generating, or Meaningful Analysis. Some of the terms that are normally used when one talk about Big data are given below.

- Data mining: It involves extracting of data from useful information from a number of different types of records including numbers, text, audio, speech, and video. Mining can be performed on both small and big data, irrespective of their size, variety, etc. Data mining algorithms, if are to be applied in big data domain would encounter issues with respect to memory capacity, memory utilization, applicability, availability, speed, flexibility and scalability [2].
- Traditional analytics and big data analytics: Traditional analytics involves limited data sets, cleansed data and simple models. It performs descriptive and diagnosis analytics. Analysis is more focussed on what happened and why it happened. It involves studying huge amount of data to reveal unknown patterns, relationships, connotations, associations and other meaningful insights. Big data analytics involves data sets of very large scale and complex data models. It involves studying huge amount of data to reveal unknown patterns, relationships, connotations, associations and other meaningful insights. It performs predictive analytics. Analysis is more focussed on new insights and more accurate answers.

Characteristics of Big Data: V-Four - C-Four (V4-C4)

The word- big data is a buzz word nowadays. The practice of collecting data and storing it for analysis is very old in the world of computing. Big data however has several important attributes [3][4] that make it different. These are now explained and shown in figure 1.

- **Volume:** Implies large amount of data ranging from terabytes to petabytes each day. As already mentioned, high volume of data is

generated from multiple sources like sensors, machines, social media, humans, automobiles etc. Storage technologies along with computing power has greatly enhanced as against the earlier days when it was tedious and expensive. Improvement in these technologies has enabled organizations to store much more data than ever before. It is expected data volume to exponentially increase 44x from 2009 to 2020; from 0.8 Zettabytes to 35 Zettabytes.

- **Velocity:** Signifies a large amount of data in small amount of time. In some situations the pace of data generation is more significant than the volume of data. For example, the speed of stock trading data, flight data, space data, and machine to machine processes data is massive and is also endlessly growing fast. In some scenarios even 5 minutes is too late, for some cases 10 nanosecond delay is too much and in some 50 hours is fine. For time-sensitive processes for example, to find out fraudulent activity, Big data must be used.
- **Variety:** Denotes unstructured, semi-structured or structured data. Much of the Big data is generated from unstructured sources including audio, video, text, social media data, and log files, whereas the structured sources involves dates, numbers, strings, sequences, time series, and multi-dimensional arrays, etc. We have static data and streaming data in single applications that are generating many types of data.
- **Variability:** Includes periodic dips and peaks in the data. Sometimes the data is highly erratic with periodic peaks and dips. For example, unstructured data can grow rapidly on social media when some event happens or is about to happen.
- **Complexity:** Entails that data that is hard to comprehend. As data is generated from multiple sources, it becomes difficult to associate, correlate, clean, pre-process the data across systems. It is important to find correlations among data for creating useful insights.
- **Cross-Domain:** Conventional data mining techniques are normally applied to data from a single domain. In the age of Big data, where diverse sets of datasets are available for analysis from multiple sources from multiple domains. These datasets consist of multitude of data with different data distribution, representation, scale and density. For this innovative machine learning based methods are deployed which can provide insights from various datasets.
- **Customization:** Big data is transforming many market segments through its ability of customization by combining datasets from various sources (internal or external). Personalized marketing, personalized banking, personalized pricing, personalized manufacturing personalized clothing are driving the organizations to build better business models and hence increased revenues [5].
- **Convergence:** Convergence of Big data and cloud technology has provided the organizations with huge benefits ranging from state of art infrastructure at minimum costs, latest software tools, flexibility of operations, ease of management and zero maintenance. Cloud technology gives the organizations one stop solution for the storage and computing power related issues. Organizations can just focus on harnessing Big data to get valuable insights.

Opportunities

Big data revolution has impacted every kind of industry segment. Some organizations have recognised the potential of Big data to solve multitude of their problems; some organizations are using Big data to help them grow by finding out core causes for failure of their systems and thereby making the systems more efficient. Some industries are still in doubt to find out if there is any benefit in deployment of Big data. There are some industries that are still in testing phase, while others are developing the products of Big data to their use. In the following section, applications of Big data in different industry segments is discussed [6][7] and shown in figure 2.

1. **Banking:** It is an obligatory requirement of the banks to lessen the risks and frauds involving in the financial transactions along with maintaining higher customer satisfaction. Due to the support of Big data analytics and advanced techniques, banks can identify the cause of risks in one system and thereby improve the efficiency of the other systems.
2. **Education:** Data is available from varied sources in the field of education. Institutions, staff and

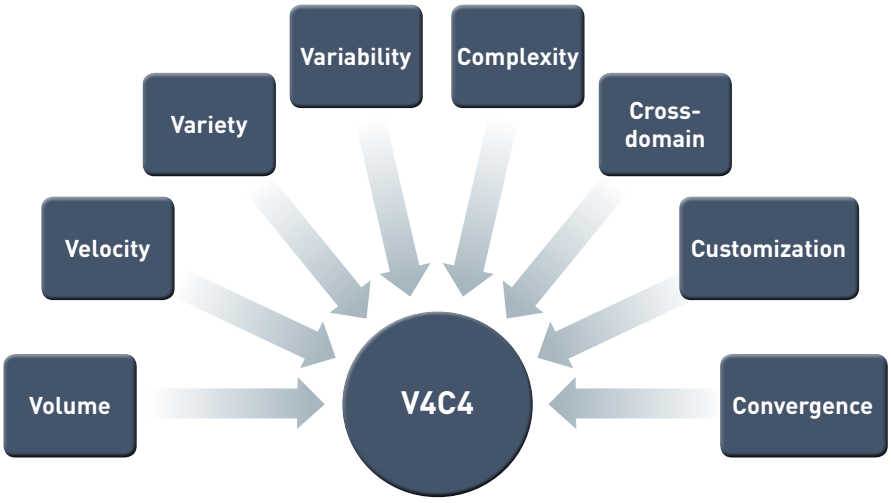


Fig. 1 : Characteristics of Big data

students are not skilled enough to use Big data. Big data can be used to measure performance of faculty, staff and students.

3. **Public domain:** In this domain, Big data has a wide range of applications including fraud detection, stock trading, weather prediction, health related research, environmental protection and market analysis.
4. **Healthcare:** Nothing can be ignored and can be taken casually in healthcare delivery organizations. These organizations demand that everything to be done in a well-organized and effective manner. A large amount of data is being generated from almost any activity in the hospital like patient health records, pharmacy, prescription information and medical reports. One can discover hidden insights from the available data for delivery of better patient care. Patterns of illness and diseases are studied using Big data to determine faster and better treatment with fewer casualties.
5. **Manufacturing:** Use of Big data analytics improve the quality of products, energy efficiency and profit margins in field of manufacturing. Increase in demand of natural resources including oils and a mineral has also led to increase of Big data in terms of complexity, velocity and volume in natural resources industries. .
6. **Retail:** Customer relationship has a great impact in the retail industry. Big data analytics helps retailers to identify the shopping patterns of the customer, the efficient ways to market their products to customers, timely analysis of their inventory, stream lined supply chain requirements, reduce fraud transactions and optimized staff placements.
7. **Recommendation Systems:** Personalized recommendation system for each user is made that helps in engaging user on the website for more duration and with much more attentiveness. Good or bad recommendation depends on how rich context model system is using for selecting information derived from analysis of Big data to display to a user.
8. **Online Advertising:** User profiling is done based on analysing his interests, content viewing pattern and stream of clicks. This helps in creating better revenue models by showing customized advertisements to a particular user.
9. **Network Monitoring:** Using Big data, root cause analysis is done for the congestion in the networks. By analysing the live feed of data, unfamiliar behaviour of the network, faults- both short and long term can be detected. .
10. **Messenger Social Network:** Using Big data techniques, one can observe consumer sentiment and their behaviour to figure out its impact on one's business and thereby making the business processes much more efficient.
11. **Transportation:** Route scheduling and planning is done in order to save on both fuel and time. Smart transport systems can be developed that can predict traffic behaviour based on various parameters and road conditions. Such kind of applications can really impact tourism industry by designing better travel itineraries and arrangements.

Fig. 2 : Application of Big Data in Different industries

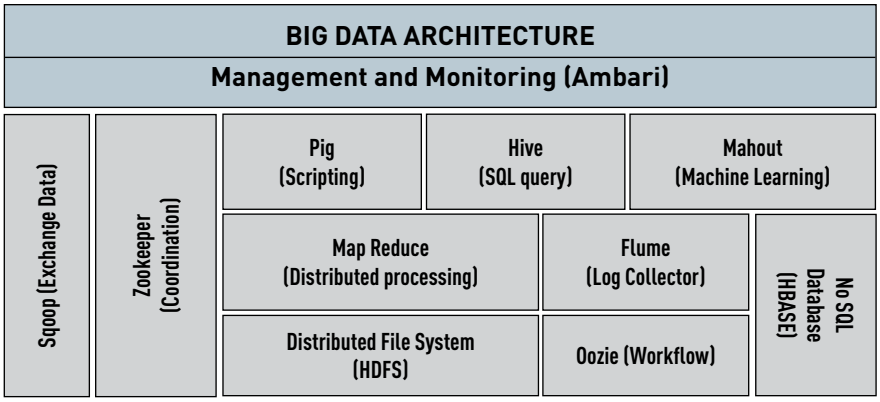


Fig. 3 : Architecture of Big Data

Architecture, Tools and Techniques

With the pace of data growth, there is development of thousands of tools and techniques to save time, money and help to uncover new insights [8] [9]. Various tools and techniques are developed for data storage and its management, cleaning, mining, visualizing and analysing. Some of the popular tools and techniques for Big data analytics are as follows and its architecture is shown in figure 3.

1. **Hadoop:** Hadoop is a software ecosystem rather than a simple database that allows support for huge parallel computing, redundancy and distributed architectures. It is a type of a distributed file system that contains data processing engine which is

able to handle exceptionally high volumes of data in any structure. There are two major components of Hadoop (i) Hadoop distributed file system (HDFS): it supports relational data ranging from structured form to semi structured form to completely unstructured form and (ii) MapReduce programing: it is a distributed processing paradigm for managing applications of different servers in distributed environment. Some terms that are closely related to Hadoop are: (i) Apache Avro: it is used for communication between nodes through data serialization, (ii) Cassandra and Hbase: it is a non-relational database to use with Hadoop, (iii) Hive: it is a query language analogous to SQL and companionable with Hadoop, (iv) Mahout: it is an artificial intelligence system used for machine learning, helps in filtering data required for examination, (v) Pig Latin: it is a data flow framework for parallel computing and (vi) ZooKeeper: it keeps all the parts synchronized and working together.

2. **NoSQL:** Not Only SQL is a database used to solve issues like (i) Big data performance and (ii) scalability. These are the challenges that relational databases cannot solve as they were not designed to address so. NoSQL helps an organization to access and analyse the massive amounts of unstructured data that is stored on virtual servers. NoSQL has moved much beyond relational database models that uses SQL and tables. It simply emphasizes on retrieving the data and attaching new data not necessarily in tabular format. It focuses on schema less architecture, key value pairs data stores that can be used to locate data objects and on supporting storage of large amounts of unstructured data. In addition, it does not have ACID (atomicity, consistency, isolation, durability) properties.
3. **MapReduce:** MapReduce is the heart of Hadoop. It is a programming model used for distributed processing of the large data. It is based on Java. MapReduce has two major components viz. Map and Reduce. First map function works on the data set followed by the Reduce function. Map takes as an input a set of data, breaks each individual element into key value pairs and in the process generates another set of data. Reduce component, on the other hand, takes the output of map (the first component) as its input, merges them and generates a smaller set of key value pairs. MapReduce can perform processing of data over numerous computing nodes. The number of these nodes can be scaled up or down depending on the application requirements. Scaling permits the application to run over numbers of thousands nodes in a cluster. This feature of scalability is one of the most important reasons of its popularity amongst the programmers.
4. **Cloud Technology:** It is a technology that enables to share a pool of storage and computing resources based on on-demand access. For storing Big data, cloud is a boon. There are many cloud technology providers in the market. For example Amazon provides Amazon EC2 (Elastic Compute Cloud), which is a web interface to provide resizable distributed servers (or clouds) for hosting the processing of Big amount of data. It allows data developers to have full control over web-scaling, storage and computing resources. EC2 instances can be resized and also their number can be decreased or increased as per the requirement. Moreover these instances can be instantiated in one or more distinct geographical locations and Availability Zones. Several Availability Zones are connected by networks with low latency in the same region.

Conclusions

In this article, the authors have examined the advanced background and roadmap for Big data. Big data has lately gained lots of attention due to its perceived distinctive opportunities. Firstly, the relation between Big data and different related technologies viz. data mining and analytics is given. Then the four V's and four C's of Big data: Volume, Velocity, Variety, Variability, Complexity, Cross-domain, Customization and Convergence are explained. Finally the article showcases several representative applications of Big data along with various tools and techniques currently being used in the industry to implement them. Big data is the technology that has an immense potential and in future is going to impact immensely in (i) improving data analytics techniques, (ii) Natural Language Processing methods to find out the current sentimental trend which has varied applications in almost each sector of business, politics, finance, (iii) designing improved data mining algorithms, (iv) data visualization techniques and (iv) merging prescriptive analytics with business analytics software's. However the road of Big data is not as simple as it looks to be. There are several challenges too that Big data will have to face; security and privacy of data being on top of the list.

References

[1] Demchenko Y, de Laat C, Membrey P. Defining architecture components of the big data ecosystem. In: Proceedings of the International Conference on Collaboration Technologies and Systems, 2014. pp 104–112.

[2] Big data across the federal government, http://www.whitehouse.gov/sites/default/files/microsites/ostp/big data fact sheet 3 29 2012.pdf, Fact sheet, 2012.

[3] Big data, http://www.nature.com/news/specials/bigdata/ index.html.

[4] Adomavicius, G. and Tuzhilin, A., Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions, IEEE Transactions on Knowledge and Data Engineering, pp. 734-749, 2005.

[5] Beyer M, Gartner says solving big data challenge involves more than just managing volumes of data, http://www.gartner.com/it/page.jsp, 2011.

[6] Bollier and David, The promise and peril of big data, Aspen Institute, 2010.

[7] Walter T, Teradata Past, Present and Future, UCI ISG lecture series on scalable data management, 2009.

[8] Big Data, www.sas.com/en_us/insights/analytics/big-data-analytics.html; 2013.

[9] Sagiroglu, S.; Sinanc, D., (20-24 May 2013), Big Data: A Review.

■

About the Authors

Dr. Sunil Gupta [CSI- F8000738], currently working as Associate Professor, Computer Science and Engineering, BML Munjal University, Gurgaon, India. He has conducted various workshop, conferences and FDP. He Guides various students for research and project work. He is an authored 55 research papers, one Patent and two books, namely Cryptography and Network Security and Wireless Sensor Networks. His academic interest includes security, cloud computing, big data, sensor and wireless networks. He can be reached at sunil.gupta@bmu.edu.in.

Prof. Goldie Gabrani [CSI-F8000737], currently working as a Professor in Computer Science and Engineering at BML Munjal University. Prior to BMU she has been a faculty member at Netaji Institute of Technology and Delhi Technological University. Her primary area of focus at the academic institutions has been to plan and develop the academic structure, curriculum, lab & research infrastructure, teaching new technologies beyond curriculum, consulting, researching, participating in conferences, publishing journals, building industry linkages etc. Her academic interests include digital hardware design, computer networks, distributed computing, data analytics, machine learning, IOT. She can be reached at goldie.gabrani@bmu.edu.in

KIND ATTENTION !
Prospective Contributors of CSI Communications

Fourth Coming Issues : December : Heritage Computing

Please note that Cover Theme for **December 2018 issue is Heritage Computing**. Articles may be submitted in the categories such as: Cover Story, Research Front, Technical Trends, Security Corner and Article. Please send your contributions by 20th November, 2018.

The articles should be authored in as original text. Plagiarism is strictly prohibited.

Please note that CSI Communications is a magazine for members at large and not a research journal for publishing full-fledged research papers. Therefore, we expect articles written at the level of general audience of varied member categories. Equations and mathematical expressions within articles are not recommended and, if absolutely necessary, should be minimum. Include a brief biography of four to six lines, indicating CSI Membership no., for each author with high resolution author photograph.

Please send your article in MS-Word format to Chief Editor, **Prof. (Dr.) S. S. Agrawal** in the email ids **csic@csi-india.org** with copies to the Publisher **Prof. A. K. Nayak**, in the email id : aknayak@iibm.in and **Prof. R. R. Deshmukh**, Prof. Babasaheb Ambedkar Marathawada University, Aurangabad, Maharashtra in the email id : ratnadeep_deshmukh@yahoo.co.in.
Issued on the behalf of Editorial Board, CSI Communications.

Prof. (Dr.) S. S. Agrawal
Chief Editor

EERTREE: An Efficient Data Structure for Processing Palindromes in Strings

Vaibhav Srivastava
Student, Computer Science, NIET, Gr. Noida, India
iamvaibhav505@gmail.com

Kalpana Dwivedi
Asst. Professor, Computer Science, NIET, Gr. Noida, India
dwivedi.dolly@gmail.com

The 'PALINDROMIC TREE' or 'eertree' is a linear size data structure that provides fast access to palindromic substring of the given string. The inventor of PALINDROMIC TREE is Mikhail Rubinchik and this paper describe basic principles of this data structure. Most of the data structure that we use are at least 40-50 years old. This data structure was presented in Petrozavodsk Summer Camp 2014 by Mikhail Rubinchik and Arseny M. Shur. After that he officially presented in an academic paper in the year 2015. This paper is mainly based on Palindromic Tree. It is fast data structure that provides fast access to the all the palindromic substrings of a string. This data structure can easily give an efficient solution to many of the palindromes related problem.

Keywords: eertree, palindromic tree, insertion edge, maximum palindromic suffix edge

I. Introduction

Palindromes are the words and phrases that read the same in both directions or we can say that palindromes are repetitive structure of strings that can be read same in both the directions. There are several papers that deals with the problems related to palindromes, some of the problems are searching and counting of palindromes in the strings. one of such most popular algorithm is Manacher algorithm, that finds all maximal substrings of string in linear time. The Eeertree or Palindromic tree consists of nodes that contain palindromic strings and two types of edges i.e. insertion edge and Maximum palindromic suffix. The time complexity of this algorithm is O(n), where n is the size of string and since the size of string is always a constant value, so we consider O(n) as a constant. This tree contains two root nodes and n nodes connected to each other using directed edges, so this tree totally contains n+2 nodes.

II. Proposed idea of EETREE

EERTREE contains n+2 nodes where n is number of palindromic substrings and two types of edges. It is like the directed graph all the nodes contain palindromic substrings of the given string for each node it is mandatory to store the length of palindromic substring.

The palindromic tree consists of following two types of edges:

1. Insertion Edge: The insertion edge or weighted edge from a node n1 to node n2 with some weight x means that a node n2 is formed by inserting x at the front and end of the string at n1, since n1 was a palindrome so the new node v will also be a palindrome. The weight for an edge connecting a palindrome 'A'(n1) with the node 'xAx'(n2) in which x is the weight. In each node may have zero or more edges outgoing from it and one incoming edge otherwise palindrome will not be reachable. In a single node we can only insert 26 words i.e. maximum of 26 insertion edge if we take lower case strings only. I am currently taking red color edges to represent insertion edge. The following diagram represents it,

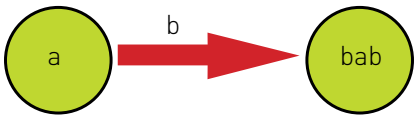


Fig. 1 : Inserting b in node n1

In the Fig. 1 we are inserting b in a (n1) and a new node bab (n2) is obtained after insertion.

2. **Maximum palindromic Suffix Edge:** This edge indicates to maximum palindromic substring of the string. I will use blue edges to represent maximum palindromic suffix edge. I will address it as suffix edge. Every node has a unique suffix edge as it will not store duplicate strings in the tree. This edge is unweighted edge. The following diagram represents it,

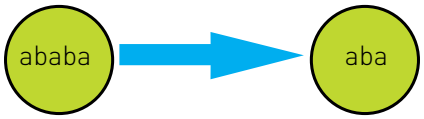


Fig. 2 : node r1 and r2

This tree consists of two root nodes let say r1 and r2. The node r1 will describe the string of length -1 and r2 node will describe a null string of length 0.

III. Implementing the palindromic tree:

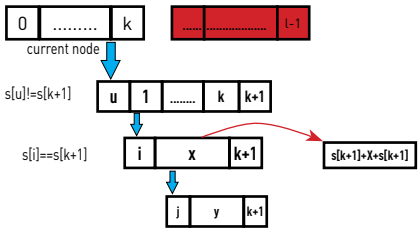
In the process of building a palindromic tree, we need to insert the character of the string one by one till we reach to the end and after we are done inserting, we have the palindromic tree.

Let's begin this process, first we take a string s of length l and we have inserted the string till k index where(k<l-1). Now we start our process by inserting the (k+1) character. so, insertion of (k+1) th character means that the palindromic string formed will be ('s[k+1]' + "X" + 's[k+1]'), here X will be a palindrome.

So from the above it is clear that to insert the character s[k+1] we only need to find X i.e. palindrome in the tree and use insertion edge and put s[k+1] as weight and direct it to a new node that will contain 's[k+1] + X + s[k+1]'.

Now it is clear that we will have to find the string X in the tree in the minimum possible time, we are storing the suffix link of all the nodes. So, to search the node that contains X we will have to traverse down the tree.

In below diagram we will find the string X and then we will link it with new node that contain s[k+1] +X+s[k+1] with weight s[k+1]. The current node in figure below that it is the largest palindrome from 0 to k. now the blue arrow at every step is the suffix link and X will be in the one of the node below it we will find it by performing multiple iterations. we will place k+1 th character at the end of every node and compare it with first character of the string, if it matches then X is found. After we get X we will connect it with the help of insertion edge to the new node with the weight s[k+1]. The Fig. 3 represent the above scenario.



Note: White boxes are processed strings and red ones are unprocessed.

Fig. 3 : processed and unprocessed strings

Let us take a string s = "abba" with length l=4.

At initial we have two nodes one has length -1 i.e. it has an imaginary string and second node have a null string of length 0, this is initial step we haven't inserted any character in the tree, the following diagram represent it,

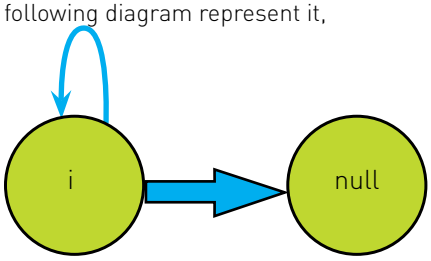


Fig. 4 : Inserting a in root node

Step 1: At first, we will insert a new character 'a' i.e. s [0], we will insert a in the root 1 having length -1 so after insertion its length becomes 1 now the string at the new node will be "a". so, will create an insertion edge from root 1 to the new node. Now the largest palindromic suffix of the string will be of length 1 and the current node will be new node that contains "a".

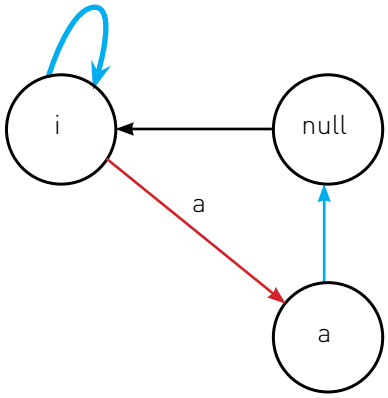
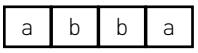


Fig. 5

Step 2: In this step we will insert 'b' i.e. s [1] in the tree. The insertion process will start from the current node, we will find X in the tree by traversing through suffix link. So, after traversing we found that the root1 is X string. Then we start insertion of 'b' to the string of length -1 that will result in a new node of length 1 and string 'b'. Now the current node will be the new node that contains string "b".



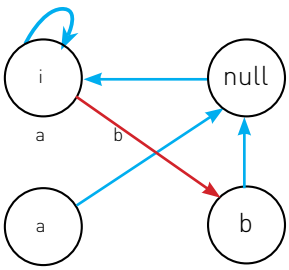
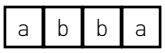


Fig. 6

Step 3: In this step we will insert 'b' i.e.s [2] in the tree.Again, we will start from the current node and will traverse down the tree using suffix link to find the required string X. The string X in this case will be root2 , so after adding b in the front and end the resulting string will be "bb" and length will be 2 .So we will construct a new node that contains "bb" by use of insertion edge that direct from null string to newly node that contains "bb" current node now becomes "bb".



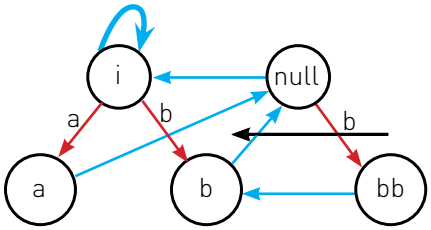
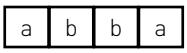


Fig. 7

Step 4: In this step we will insert 'a' i.e. s [3] in the tree, after traversing we will find that largest string X will be current node such that s [0] + X + s[3] is the palindrome.So, we will create a new node that contains "abba" as the string. We will create a new node by using insertion edge having weight 'a' from current node to new node that contains "abba".



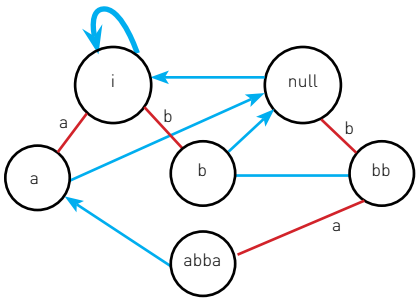


Fig. 8

Time complexity:

The time complexity of this algorithm is O(n), where n is the size of string and since the size of string is always a constant value, so we consider O(n) as a constant.

III. Conclusion

In this paper, we tried to explain the basic principles and concepts of 'Palindromic tree' or 'eertree', which stores all the palindromic substrings of the given string. This data structure provides fast access to all the substring of the given string.

IV. Future Work

This topic has wide scope in the field of problems related to palindromic strings with the help of this data structure one can easily solve future problems related to strings in linear time and problem related to the optimal construction of eertree.

Acknowledgment

I wish a very thank tour colleagues from NIET who provided insight and expertise that greatly assisted the paper.

I would also like to show my gratitude to Ms Kalpana Dwivedi (Asst. Professor (CSE), NIET, Gr. Noida) for sharing their pearls of wisdom with meduring this research, and I also thank to the reviewers for their so-called insights.

References

[1] http://adilet.org/blog/25-09-14/

[2] https://www.geeksforgeeks.org/palindromic-tree-introduction-implementation/

[3] https://medium.com/@alessiopiergiacomi/eertree-or-palindromic-tree-82453e75025b

[4] https://pdfs.semanticscholar.org/6087/f2ff4f58df199c209655b0244597e72792b6.pdf

■

About the Authors

Vaibhav Srivastava is a student Of Computer Science& Engineering from NIET, Gr. Noida and he is writing his first research paper in Data Structure by using Palindromic Tree as topic.Enriched with ability to quickly understand & implement new technologies

Area of Interest: PEGA Certified System Architect, JAVA, C. He is pursuing his Bachelors of Technology from Noida Institute of Engineering & Technology.

Kalpana Dwivedi is a Asst. Professor of Computer Science Engineering in NIET, Gr. Noida. She has 5 years of Experience in Teaching and research work.

Area of Interest: Big Data Technology like Hadoop, Splunk, Mongo DB, Cassandra
Her academic qualification is MTech (CSE), BTech (CSE). Her research paper is published in IEEE.

About the Guest Editor

Subhash Chandra Yadav

At present working as Associate Professor and Head,department of Compute Science & Technology, Central University of Jharkhand, Brambe, Ranchi. Bearing additional responsibilty of Incharge, University Technical Cell and Estate Officer (Incharge).

Having 21 years of teaching and IT industry experience Dr. Yadav is a young and dynamic academician in the field of Computer Science and Technology. Apart from carrying academic responsibility he has contributed to CSI as a founder chairman (Varanasi Chapter) and to ISCA in the capacity of Sectional Member, Sectional Recorder, and also Sectional President.

With the educational degrees like PhD, M. Phil (CS), M. Tech(IT), MCA, M. Sc and MBA (IT&HR) his area of interest are Network Security, Computer Databases, Cloud Computing, Soft Computing. Dr Yadav has received several awards and academic distinctions of national repute like Platinum Jubilee Speaker, 105-ISCC-2018, Imphal (Manipur), Significant Contribution Award for the year 2010, Computer Society of India, at IIM Ahmedabad on 1-12-2011, Purvanchal Samman for the year 2012, for significant contribution in education in eastern UP by Vijay Dubey Foundation, Varanasi, Significant Contribution Award for the year 2011, Computer Society of India, at University of Kolkata on 1-12-2012.

Dr. Yadav has authored 4 books, 5 book chapters and 26 papers of the national and international repute.

Persistent Speech Recognition using Deep Learning

Dhrmendra Patel
Associated Professor, Faculty of Computer Science and Applications at Charusat, Changa

Atul Patel
Dean and Professor of Faculty of Computer Science and Applications at Charusat, Changa

Audio data are essential data in the contemporary context of the computer science. Audio files are stored in the computer in compressed form. The recognition of audio in quality form is the challenging task. Speech Recognition research has many dimensions. In this article we will talk about persistent speech recognition dimension. Hidden Markov Model and Stochastic Processing are the vital methods for persistent speech recognition. However the precision of the yield isn't up to the mark.Deep Learning based strategies are the most recent progression in the field of persistent speech recognition. In this article, we have expounded famous strategies for Deep Learning in the setting to persistent speech recognition.

Key Words : Deep Learning, Speech Retrieval, Hidden Markov Model, Deep Neural Network, Recurrent versions of Neural Network, Convolutional Neural Network

1. Introduction

Speech retrieval is the exceptional type of the information retrieval where information is in the audio form. Speech retrieval system ought to have the capability to store,retrieve and present the substantial audio data dynamically, based on user query. The traditional text retrieval algorithms might be applied to speech retrieval easily if we could generate the transcripts of speech documents correctly. Automatic quality speech transcript generation is the advanced research area of this domain. Basically, there are three main areas of automatic speech retrieval research depicted as a Fig-1.

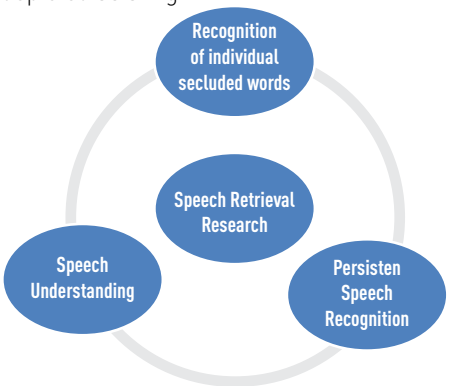


Fig. 1 : Automatic Speech Retrieval Research Areas

Recognition of individual secluded words has no such a great amount of significance as there is no genuine use of such assignment. Persistent Speech Recognition is the principle region of research in the speech retrieval space. Speech understanding is the fundamental piece of both sort of research as it manages the comprehension of the speech effectively by the end client (Lalit R. Bahl at.al.,1983). This article is confined to persistent speech recognition. The task domains of persistent speech recognition are of two types : (a) simulated undertakings and (b) natural assignments. In simulated undertakings , experimenter characterizes the grammar priori. In natural tasks, the model is to to be designed from the observed data.

Fig. 2 : Acoustic Processor

The Persistent Speech Recognition system requires the infrastructure like acoustic processor and linguistic decoder.

Fig. 3 : Decoder

There are several common research tools are available such as Sphinx, SRILM, HTK and so forth that have been utilized broadly by many research labs of speech recognition.

Statistics play a very important role in the designing of an appropriate algorithm related to speech recognition. Hidden Markov Model and Stochastic Processing are vital statistics used in the designing of speech recognition related algorithms (J.K.Baker,1975). Decision Tree based algorithms are widely used for categorization of features in the context of speech

recognition(L.Breiman et. al. 1984). Speech recognition has a long history from the algorithmic viewpoint. Most recently, the field has benefited from progresses in deep learning. The overall business appropriation of an assortment of deep learning strategies in outlining and conveying speech recognition frameworks have been expanded a considerable measure.

2. Deep Learning

Deep Learning is the new aspect of machine learning(Bengio Y,2013). It involves deep artificial neural network based algorithms to deal with machine learning problems. Deep learning based neural network comprises of neural network with numerous layers. Layers of neural network contain the weights. In the training phase, it tries to locate the appropriate measure of weight that prompts most exact yield. The Cost function is the essential measure that we need to choose while building a neural network. Optimization techniques are utilized as a part in such a a way that it minimizes the cost function.

Deep Learning has revolutionized the machine learning with some propelled techniques which have drastically enhanced the best in class in speech recognition. Completely associated Deep Neural Network(DNN) is the most encouraging strategy of Deep Learning that gives essentially higher exactness in persistent speech recognition in the correlation of prior procedures (Li Deng et. al.,2014). RNN (Recurrent versions of Neural Network) and CNN(Convolutional Neural Network) are the current progression of Deep Learning in the context to the speech recognition.

3. Deep Learning Based Methods For Persistent Speech Recognition

Deep Neural Network (DNN) is the key method of Deep Learning for the persistent speech recognition. The deep neural network has more depth in the comparison of common single hidden layer neural network. Deep Neural Network has more than three layers. Each layer of nodes trains on a distinct set of features based on the output of the previous layer. The further you advance into the neural net, the more complex the features your nodes can recognize, since they aggregate and recombine features from the previous layer. Deep-learning networks perform automatic feature extraction without human intervention. As far as speech recognition is concerned, DNN works on three main parameters :

(a) Interconnect patterns among different layers of nodes.
(b) Training process for the modification of weight.
(c) Activation function responsible for the conversion of the weighted input to output.

The flow of speech recognition using DNN is depicted in the following figure :

Feature extraction implies the identification of a speech signal that are good for the linguistic content. Mel Frequency Cepstral Coefficients(MFCCs) are standard features utilized in persistent speech recognition. They are based on frequency domains using the Mel scale. Mel scale is based on the human ear scale. Body linear predictive codes(LPC) are another standard for the features. For the medium or low bit rate coder, LPC is widely used. Perceptual Linear Prediction is another important technique for the features. It is based on the psychophysics of hearing.

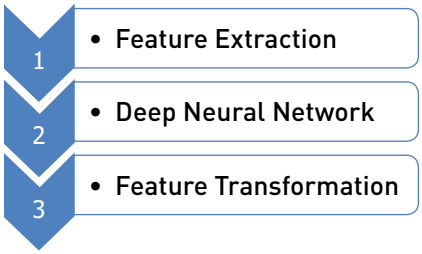


Fig. 4 : DNN based Speech Recognition Process Flow

Deep Neural Network with Hidden Markov Model has achieved incredible success on persistent speech recognition. Deep Neural Network is used to produce posterior probabilities over Hidden Markov Model directly. It also allows discrete and continuous features of speech recognition. Recurrent Neural Network(RNN) is the advanced version of Deep Learning and it is best suitable for speech recognition because of their modeling capabilities. RNN recognizes the speech as a time dependent problem. RNN reads the speech from left-to-right in time. The

Table-1 : Characteristics of Deep Learning based Methods

Sr. No.	Method	Characteristics	Disadvantage(s)
1.	DNN HMM	▪ Classical Method ▪ Good quality	▪ Data Fragmentation Problem ▪ Limited capabilities of Gaussian distributions
2.	RNN	▪ Advanced Method ▪ Speech as a time dependent problem	▪ Read the speech is only in one direction ▪ Not capable of Parallel Processing
3.	RNN LSTM	▪ Advanced Method ▪ Speech as a time dependent problem ▪ Reading the speech in both directions	▪ Not capable of Parallel Processing
4.	CNN	▪ Advanced Method ▪ Massive parallelism can be achieved	▪ Not flexible
5.	TCN	▪ Latest Advancement of CNN ▪ Greater Accuracy ▪ More Flexible	▪ Complex Structure

major improvements to RNN are done by introducing two Long Short Term Memory (LSTM) encoders that are capable of reading the speech in both directions. The major real problem of RNN is that it has not capability of parallel processing as it reads and interprets the speech at a time. The solution of the above problem is the Convolutional Neural Nets (CNNs). Massive parallelism can be achieved using CNN. CNN performs the processing by allowing number of layers between input and output. Temporal Convolutional Nets(TCNs) are the latest advancement of the CNN that produces the greater accuracy and the excellent speed of the processing. The other main advantage is more flexibility in terms of memory size controlling. The following table depicts the characteristics of the Deep Learning based methods.

4. Conclusions

Deep Learning is the emerging area of machine learning that can be applied to persistent speech recognition extensively. Deep Learning with Hidden Markov Model is the popular method of Persistent speech recognition. However, it suffers from the limitations of data fragmentation and Gaussian distribution. Recurrent Neural Network(RNN) is the ideal for the speech recognition due to their modeling capabilities. However , the technique is slow as not capable for the parallel processing. The solution of the above problem is the Convolutional Neural Nets (CNNs). Massive parallelism can be achieved using CNN. The latest version of the CNN is the Temporal Convolutional Nets(TCNs) that is more adaptable, quick and accomplishes the more prominent accuracy in the context to the persistent speech recognition. We inferred that if Deep Learning is connected with the proper strategy to determine persistent speech recognition, the mesh results could be obtained with less time period.

References

[1] Lalit R. Bahl et.al (1983). A Maximum Likelihood Approach to Continuous Speech Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-5, NO. 2, pp. 179-190.

[2] J. K. Baker (1975). Stochastic Modeling for Automatic Speech Recognition. in Speech Recognition, edited by D. R. Reddy, Academic Press.

[3] L. Breiman et. al. (1984).Classification and Regression Trees. Wadsworth & Brooks, Pacific Grove.

[4] Bengio Y (2013). Deep learning of representations: looking forward. in: Statistical Language and Speech Processing, pp. 1--37, Springer.

[5] https://towardsdatascience.com/intro-to-deep-learning-d5caceedcf85.

[6] Li Deng et. al.(2014). Ensemble Deep Learning for Speech Recognition. 15th Annual conference of the international speech communication association, Siganpore.

About the Authors

Dr. Dhrmendra Patel (CSI Membership No. 2060000012) received his Master of Computer Application degree from North Gujrat University. He received his Ph.D degree in computer science from Kadi Sarva Viswavidyalaya. His area of research is Web Mining, Fog Computing, Image Processing, Interenet of Things etc. He has published 20 papers in national/international journal of repute. Currently he is working as an associate professor at CHARUSAT, Changa.

Dr. Atul Patel (Life Membership No. 00114194) is working as a Dean and Professor of Faculty of Computer Science and Applications at CHARUSAT, Changa. He received his Ph.D Degree in computer science from Sardar Patel University. His area of research is Wireless Communications, Network Security, Cloud Computing, Internet of Things etc. He has published more than 30 research papers in national/international journal of repute.

Role of Big Data in Agriculture

S. N. Kakarwal
Prof. in Dept. of CS & Engg. of
P.E.S. College of Engineering, Aurangabad

D. V. Kurmude
Associate Professor and Head in Dept. of Physics
at Milind College of Science, Aurangabad

R. R. Deshmukh
Professor in Dept. of CSIT,
Dr. B.A.M. University, Aurangabad

1. Introduction

Information is important part of today's world. Due to availability of high speed internet connectivity, user can access the information on their laptops, desktops and on smart phones at any time and at any location. Large volume of data is created and pass on from one location to other location using internet.

Businesses are also working on such large volume of data to process it, analyse it for their product efficiency. The data on which businesses are working is also available in structured and unstructured format. It is hard to process such data using traditional data processing technique. The success of any product of business is totally dependent on market analysis of that product. This analysis totally depends on large amount of data. To work on such large volume and variety of data, Businesses need to work on big data concept. This big data concept is essential for overcome the problem of traditional approaches.

2. Big data analytics in various industry

The businesses which uses the big data analytics to grow their business are like education field, banking sectors, Health care sectors and retail sectors.

The few thrust area where Big data can be beneficial for finding the answer to the questions which are not solved using traditional techniques such as food chain market, agricultural sectors, traffic controlling, driverless car projects and many mores.

This big data can be beneficial for food chain market. It is beneficial for agricultural sector.

This sector is unstructured one where farmers are deciding to grow their farming independently.

Big data can be helpful to such farmers to decide which farming is useful for which period.

Based on market analysis and environmental analysis, farmer can decide to select the crop.

Using Big data, government can take the decision based on analysis of big data.

Government can go for the agricultural survey which is beneficial for finding the drought affected regions as well as flood affected regions. Using crop growing area and affected area, government agencies can able to find out the total loss and package given to that affected farmer. This will beneficial to improve the process of giving packages to really affected people.

For insurance companies in agricultural field, they can also use big data tool to find the land area which is insured by farmer and actual affected land area and level of affected area.

Big data can be useful for traffic control system by taking online feedbacks.

In automobile sector, Big data is useful for automated driverless car project where driverless car producing the signal from their sensors and these signal data can be captured for analysis purpose using big data tools. Information which can be analysis are road side view, climate condition for visibility purpose, accidental prone zones.

For analysis of heavy traffic area, Big Data can play important role to divert the traffic through another route during that timing.

In above all condition, Big data tools can give the better solution in minimum amount of time for scattered data as compare to traditional software's.

The need of accurate and faster result for the large number of populations can be satisfied using big data analytics tools[1][2].

3. Big data in Agriculture Field

Agriculture sector growth depends upon the weather, productivity of soil, seeds quality, pesticides, temperature require for crops, prices of farming product decided by government, transportation facilities availability, cold storage capacity availability for crops.

All above parameters are highly volatile and unpredictable in nature by common man. Due to lack of knowledge of use of technology to understand the need of market, many farmers take the wrong decision during all development stages of crops. When crop is ready to go for market that time also farmers are not getting accurate information about current market price of their product which leads -to sales their product in moderate rate decided by agents.

The complete cyclic process from seed sowing to crop marketing is highly dependent on above mention parameters. This situation can be overcome by use of big data technology.

Satsure Pvt. Ltd is using big data analytics along with Machine learning tools for soil related issues and crop process. With the help of satellite imaging they are also helping the insurance companies to understand the loss affected areas.

Using latest technologies tools with big data, SatSure Limited is an IT firm provide accurate result to bank, insurance companies and farmers to increase their productivity, reduce their risk factor of crop loss. They are working on farmers issues to resolve their problem by using big data analytics with the help of satellite imaging. They also provided the prediction about next year'sharvesting. They coin the new term as smart farming by using data analytics tools along with machine learning and Internet of Things.

This concept is fruitful for farmers to be part of new era of farming by using cutting age technology. These technologies bypass the all traditional techniques through which it is impossible to process on huge dataset.

In agricultural sector, to reduce the risk factors, it is needed to gather the data from last few years related to crop

production, weather reports, rainfall seasons statistics, satellite images of lands, insurance companies records, bank loan details, pesticides used details, temperature details and so on.

This huge information can not be process using traditional data mining technique, so big data is useful to work on such large dataset. Using big analytics tool such as MapReduce and Yarn data can be processing to get the analysis report. This data analysis report is beneficial for prediction of climate situation, soil fertility, region wise crop selection for the season.

India's economical growth is depends on agriculture sectors as most of peoples (approximately 40%) in India depends on this sector [1][2].

As per the article publish online by Wipro insights on Big data for agriculture sector, new term is introduce as Big data farming also called as precise farming. As per their report to feed large number of population of world it is necessary to have high productive crops. This can be achieved by accurate judgment of weather forecast, temperature control system and soil fertility report. This thing can be achievable by using big data analytics with Artificial intelligence.

Approximate prediction of market prices of food grains at region wise as well as international market will add benefit directly towards farmers pocket. The farmer can take their decision when to go for market, which market will give better money for their grains, the timespan of highest value for their product. The farmers are going to be self-dependent for taking decision for marketing their product. This can be achieved by using big analytical tools in farming. USA have already started such prediction system through which farmers get next seven-year forecasting of weather.

The precision farming is improving the result not only in the crop growing phase but also in crop selling to market towards end customer phase. It removes the intermediate system which reduces the time from hours to minutes [3].

As per the article of Hindustan Time newspaper, NITI Ayog has took the decision to implement a big data along with AI tools for better result in agricultural sector. In their pilot project 10 district of different states having different geographical location has selected to undergone big data analytics. The aim of this project is to increase the crop growing capacity by giving accurate prediction with the help of big data tools. Bosch have already designed their product which predicted the climate control, irrigation system controlling. Bosch irrigation system is fully automatic provide smart irrigation for water management. This company also design the automated system for polyhouse project where small change in temperature, moisture of soil can be notified by SMS service [4].

4. Conclusion

The agricultural sector is volatile in nature due to dependencies on various parameter like weather forecasting, temperature control, soil fertility, supply chain. The use of big data analytics along with live data as well as past years historical data can be useful to predict accurate weather forecasting, temperature controlling, soil nature. This will beneficial for farmers to become self-dependent for taking decision for seed growing, pesticide controlling, irrigation controlling and monitoring the market prices for their crop yield. This will reduce the loss of farmers as well as bank, insurance companies, government sector agencies who act as stakeholders for agriculture sector.

References

[1] https://www.analyticsvidhya.com/blog/2018/05/data-analytics-in-the-indian-agriculture-industry/ Dated 18-10-2018.

[2] K. Ravisankar, K. Sidhardha, Prabadevi B, Analysis of Agricultural Data Using Big Data Analytics, Journal of Chemical and Pharmaceutical Sciences, JCPS Volume 10 Issue 3, July-September 2017,1132-1134.

[3] https://www.wipro.com/blogs/wipro-insights/big-data--revolutionizing-agriculture/ Dated 18-10-2018.

[4] https://www.hindustantimes.com/india-news/govt-aims-to-harness-big-data-ai-in-agriculture-sector/story-aSOKHuPHcozPH6oFn4TC1O.html dated 18-10-2018. ■

About the Authors

Dr. S. N. Kakarwal [CSI-F8000602], presently working as Professor in Department of Computer Science & Engineering of P.E.S. College of Engineering, Aurangabad, MS-India. Her research interests include Image Processing, Pattern Recognition. She has published 17 papers in International Journals, 28 papers in International Conference and 12 papers in National Conferences. She has published 4 articles in CSI magazine. She can be reached at s_kakarwal@yahoo.com

Dr. D. V. Kurmude [CSI-2010001037] presently working as Associate Professor and Head in Department of Physics at Milind College of Science, Aurangabad. MS-India. He has published more than 20 research papers in International Journals and conferences. Authored one book published by international publisher.

Dr. R. R. Deshmukh [CSI- 00100518], currently working as Professor in Dept. of CSIT, Dr. B.A.M. University, Aurangabad, (MS), India. He has been elected as sectional member of ICT section of Indian Science congress Association. His areas of specialization are Human Computer Interaction, Digital Speech Signal processing, Computational Auditory Scene Analysis (CASA), Neural Networks etc. He can be reached at ratnadeep_deshmukh@yahoo.co.in.

Big Data Analysis – Lifeline of Modern World

Aditya Bhatia
Houston, USA

Sanjay Bhatia
ERP professional based in Houston, USA

Complex Algorithms – Simple use even for a Teenager

I turned 15 years old in Feb this year and knew that this is a milestone age in United States which bring you some special privileges. I heard unconfirmed talks within my friend circle that being 15 years old in US get you a car Driving Permit, travel alone on commercial flights and many more. Surfing internet next day of my birthday for few minutes brought me all the information I needed to know about my newly acquired status. Thanks to Big Data and algorithms behind these, I became a very informed young student about my new rights.

And thanks to Big data again, I was smartly able to use these in next few months to travel half the planet on my own, get car driving permit and research on getting funding for my Apps/Start Ups. More importantly, I can plan my path to my dream college of Harvard Business School and know what I need to do in next couple of years to get there. All I need is an average Laptop, a Smartphone, and decent Internet Connection.

Big Data – Connecting People

On completion of 9th grade in Houston, Texas, I got a 2-month vacation and wanted to visit India to meet my grandparents, uncles, and cousins. Due to professional commitments, my parents could not take time out to visit India from USA. I was also in an Immigration transition time from visa to Green Card and hence need to carefully analyze risks of returning to USA. The air travel from USA to India involve landings in some European or Middle East countries with some risk of transit passes etc. In summary, I needed to do lots of research before starting my travel.

Internet and Big data brought all this information on my fingertips within few minutes. I could study pretty much everything on airline, immigration and transit visa rules of various countries and choose the best route to travel. Many travel portals even helped me book the most economical air tickets, shuttles and even meals of my choice on airlines. The day of my travel happened to be of heavy rain in Mumbai with almost entire city coming to the halt due to water logging. However constant availability of data helped me take smart decisions all along the route.

Thanks to Big Data, I was able to travel to India safely, spend few weeks with my family and return to US safely.

What is Big Data Analytics?

Big data analytics? What is big data analytics? Even a teenager like me learning advanced computer science did not know what big data analysis meant till few days back even though unconsciously using it day in and day out for past few months. However, a couple of minutes on the internet I was quickly able to understand what this term meant, write this article, and get it published in the prestigious CSI publication that you are reading right now.

90% of all the data you see on the internet was created in only the last 3 years. All this data makes up what is known as the internet which can be accessed through your smartphone, smartwatch, laptop, PC, or any smart device.

This data consists of items such as News Articles, credit card transactions, bank transfers, search engine searches, text messages, emails, online videos, and virtually everything put on the internet.

Now, try to understand that having so much data is very difficult to sort through. Imagine you were given the task of monitoring every financial transaction in the world and had to flag fraudulent transactions. With almost 10,000 transactions every second, how would you be able to do it?

This is a job for big data analytics. This is what Google and other search engines do. Its job is to search the entire worlds information and bring up exactly what you are looking for, all in a matter of seconds.

How search engines work is a topic for another day, but the key idea is that big data analytics is taking large amounts of data, analyzing, and processing them to return a result.

Airline Tickets -

A simple example of this is airline search engines. For my recent trip to Mumbai earlier in June, my parents wanted me to search for air tickets. I used Google Flights for this purpose. Google Flights tracks and sells tickets on almost all passenger flights in the world.

Every single day there are thousands of individual flights, about 102,000 as per estimates. Here is the worldwide air traffic as I write this article

Each of these planes will have around 100-250 seats and some will have as many as 600 seats, meaning in just 1 day there are tens of millions of seats to be sold by the airlines worldwide. And Google Flights allows you to buy air tickets up to 12 months in advance, meaning there are more than 5 billion airline seats that Google Flights must keep track of, along with millions of individual flights.

Yet when I enter my dates and budget it is able to return results to me in seconds.

Many industries rely on these systems, not just flight ticket engines. The key to analyzing such giant heaps of data lies in the 5 V's

5V's of Big Data

Volume – the size of the data
Variety – the different kinds of data

Velocity – the speed of data processing

Variability – the consistency of the data

Veracity – the quality of the data

The massive amounts of data are plugged into specially developed algebraic sequences whose job is analyzing and filtering the data that they receive. Thus, when I ask for tickets to India on a specific date, the analytical algorithms are able to process my commands to deliver the result I need them to. The rest of the data not returned will be sent back where it came from until it is called upon.

Such algebraic methods are used in much more heavy ways such as monitoring stocks for investment banks such as Goldman Sachs and Merrill Lynch. These banks have trillions of American dollars in assets in the stock market which they need large algorithms to constantly keep track of.

Big data analysis is critical to keeping the stock market, search engines, airlines, and almost every industry in the world up and running. And hence the title of this article – It's a lifeline for Modern world. ■

About the Authors

Mr. Aditya Bhatia (Son of CSI member Sanjay Bhatia, CSI – 1161672) is a 15-year old student of 10 grade living in Houston, USA. He has already launched mobile app "SportsConnect" at the age of 13 which has few thousand downloads. Aditya holds 3 software copyrights, 2 Trademarks and has filed 2 Patents so far. He has been featured in 21 newspapers in USA and India about his work in the field of technology. Aditya is working on new app launch "MeriPheri" along with his teenager cousin in India and actively exploring angel investment for his technology ventures. He also dreams to study at Harvard Business School.

Mr. Sanjay Bhatia (CSI – 1161672) is a ERP professional based in Houston, USA. Sanjay has advised many Fortune 15 companies on their Real Estate Portfolio Management on ERP. Sanjay holds 11 software copyrights, 3 Trademarks and filed 4 patents. His ERP based data migration products are used by corporates in USA and Australia. Sanjay is actively mentoring Aditya to chase his dreams of startups and technology ventures.

Big Data in Business Intelligence

S. Balakrishnan
Professor, Department of Information Technology, Sri Krishna College of Engineering and Technology, Coimbatore, Tamilnadu, India.

Rahul R.
Final Year, Department of Information Technology, Sri Krishna College of Engg. and Tech., Coimbatore, Tamilnadu, India.

Overview

As the world is now evolving customer-centric, and all smaller ideas and techniques are made into business models the world has become the space for the corporate race. All companies now-a-days performing to maximum extents and striving to stay forward from their competent, they do different things and do things differently! though they all have very fast thinking minds behind them, it is hard for a company to stay in that leading position for a prolonged time, and obviously the co-runner comes before. And again, the company rushes leaving its defender back. Thus, this cycle goes on! The most different thing the companies do is the Business Intelligence. The difference in success between the companies in the current situation is just a thread-ticked. To achieve success through that atom sized cranny, most of the companies apply the most powerful technology called Business Intelligence and let us look deep into the warehoused information about the warehousing concept itself in this article.

Introduction to Business Intelligence

Business intelligence (BI) is an ideology of rendering out the right information or desicion to the people requiring them at the right time, to improve business situations and performance. It is an umbrella concept, which includes various softwares, data structures, big data concepts, algorithms etc.,

It is like the other data interpretation concepts, which requires data as it's main object and undergoes usual processes such as data collection, storage, processing and extraction of information too. The information retained or desicions thus retained after the pre-processing are represented in the froms of reports, graphs, summaries, charts etc.,

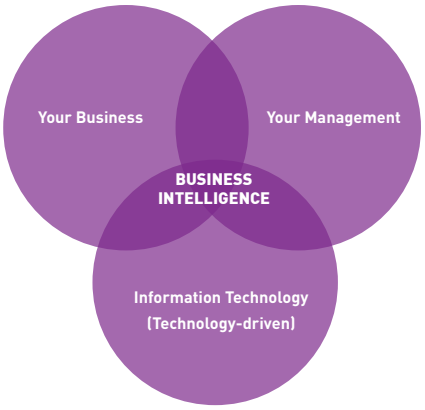
The results of BI have a direct effect on company's business operations, working strategies and tactics as the companies and organisations change their operating methods accordingly and adopt different strategies based upon the decisions of BI. So, it is mandatory for the decisions to be reliable and accurate as they are what going to decide the future of the company and helps the management to act before their competents.

History of Business Intelligence

The concept has been in existance all times through-out, right from a native shop-keeper remembering his frequent customers and the product they ask-for to a larger company producing a range of products that is in demand in the market! The concept has been realized at first by Richard Millar Devens when he represented the term 'Business Intelligence' in the Cyclopedia of Commercial and Business Anectodes. This was in 1865.

Later, the term was again remembered when a computer scientist at IBM named Hans Peter Luhn, as he mentioned it in his article in 1958. Now the concept is vibrant due it's effect and efficiency in business upliftment. For example, from Luhn's explanation about a banker Sir Henry Furnese, obtained profit by collecting information and acting before his competents.

It is still easy to understand business intelligence by knowing the Webster Noah's definition stating



that "the ability to apprehend the interrelationships of presented facts in such a way as to guide action towards a desired goal". It is evident that BI has been evolved from Decision Support Systems(DSS), which itself evolved from the computer aided models for assisting desicion making and planning.

Role of Big Data in Business Intelligence

The cardinal element of business intelligence is data. Big Data points on the volume of both structured and unstructured data collected from the sources. The size of data relies upon the sources of data considered, the company's establishment in the market, it's short and long-term goals to be achieved, knowing it's customers' need, it's business model etc., It is simple to explain the big data in just three words, variety, velocity and volume of data.

Big data invloves in these main activities with data i.e: collection, storage, integration.

Collection:

The techniques of collecting the data by the companies have transformed versatilely. The feedback method has now become an old technique. Now a day the customer requirements are learnt just by observing them and not even taking the method to their knowledge. Those few tricks are, by

using high sensitive cameras with accurate motion sensing, by tracking the online purchasing method of user, by noticing the online transaction data etc.,

Storage:

As the amount of data collected have become vulptous, the warehousing facility should also be bulged up. As the size of big data is said to range between few dozen terabytes to exabytes, the storage area may be requiered with double of their sizes predicted. Few methods of storing big data are Hadoop and Mapreduce (they usaually do analysis with SAS, Splunk and SAP Hana), Edge Computing (the simultaneously generating data has to be stored in an expanding storage space), Multi-Cloud (the public cloud platforms offering online computing opportunities), Storage Intelligence (this the software's themselves harnessing the requiered storage space).

Integration:

This is the joining the closely relevant processed data together, based on their relativity. This stage usually comes across various challenges to be faced. Few are, the information extracted might haven't magaed properly, finding the right place to place the data in big data, synchronisation of data sources, talent lack in handling the data while incorporating them in the technologies and other un-expected miscellaneous challenges.

Techniques of Decision Making

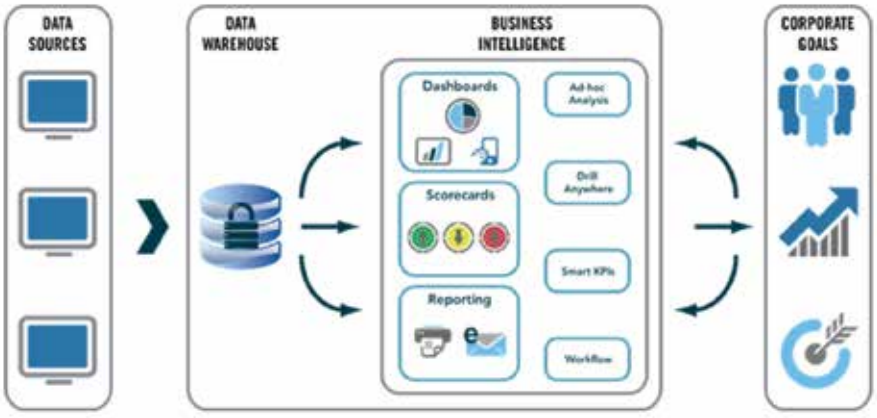
These are few latest algorithms applied for decision making.

i) Hypothesis Testing:

It is a simple idea to analyse to check whether the hypothesis is true or false based upon the data given. With the results of this hypothetical testing, we could either accept or reject the hypothesis. This method is useful in analysing whether an event happend is either by chance or an upcoming trend.

It is of two types; T-Test, a statistical test based upon the comparison of two different means or variances, how far they are different from each other. And the other one is Chi-Square Test, to test variability of two distributions of variables.

The test formula is,



$X^2 = \Sigma(0-E)^2/E$

Where,

X^2X^2Chi Square obtained

O=observed score

E=expected score

ii) Random Forest:

Decision trees in collection is called a Random Forest. It is used to classify the new objects based on it's attributes, thus with that each tree is classified. The forest chooses the classification having the most number of votes. For a better understanding, let us consider a dataset with N sample cases, every case is chosen at random and this set is the training set for growing the tree.

If the number of cases is M, all cases are tested without pruning and the tree is grown to the largest extent possible.

iii) Gradient Boosting & AdaBoost:

When predictions with higher accuracy must be made eventhough the data load are massive it is suitable to go with these algorithms. It is an ensemble learning algorithm that combines the predictive power of several base evaluators to improve robustness.

These boosting algorithms are employed in data science competitions like AV Hackathon, Kaggle, CrowdAnalytics. To achieve high rate of accuracy it is being adviced touse along with Python and R. These could also be partnered with other multiple weak or average predictors to build a better predictor.

iv) Market Basket Analysis (Apriori algorithm):

Apriori algorithm, to draw data insights on which products are likely to be brought together and the things which are most responsive to promotion. For example, a shopkeeper may apply this idea to predict that the people who buy tea and milk are likely to buy sugar and lemon to make tea. This ideology is applied by giants like Amazon, Walmart etc.,

v) Logistic Regression:

It is the method of statistically analyzing a dataset in which one or more independent attributes that determine the outcome. It predicts the outcome that could only have two values, i.e: dichotomy. For example, in predicting win or lose type of events.

$$Z=\beta_0+\beta_1 \times X_1+\beta_2 \times X_2$$

Tools

Algorithms in the form of a software is itself a biggest business as they are themselves the core functionality of a coded product. Few of those software buzzing in the machines and looting away huge sum are below:

i) Tableau:

The fifteen-year-old Seattle company is a forerunning business analytics. It is considered as the most advanced business software as it connects to various sources of data with it's productivity suite. It is said to be the most technological and high speed as it offers 10 to 100 times faster analytics than its competitors.

ii) Power BI:

The Microsoft's product is the other leadingly adopted analytics software. It offers cloud-based analytics facility too through which the users could build re-usable data models with the existing data. It also gives self-servicing access to the other major third-party cloud

sources such as GitHub, Zendesk, Marketo, Salesforce etc.,

iii) Cognos:

The oldest ever business intelligence company in the field since 1969 by IBM. It initiated its services as a consulting company for Canadian federal government. Later the company focused on software sales leaving consulting services. Further it re-newed itself by developing its own web technologies as its competitors were adopting new technologies then. Now it is producing software compatible to both mobile and desktop devices.

iv) Qlik Sense:

The company from Pennsylvania, US is other efficient software in current days. The spaciality of this tool is, it's capability to integrate automatically the visualizations and respond changes in context from anywhere in the software. It provides user a responsive user ezperience and intelligently adapts al the fuctionalities requiered from any sort of devices with touch.

v) SiSense:

It holds the number one ranking in the market of business intelligence. This is the first company in the field

to provide services with in-memory technology called in-chip analytics. This was done to maximise disk space, memory and CPU performance. It also launched Prism 10X, which helps in processing the data with 10 more efficiency.

Future of Business Intelligence

Embedded Intelligence:

As everything has its own future, let us look what BI has got to it. It is said to be embedded intelligence. In this softwares are said to have the capabilities of reporting and analytical functionalities. This feature may be present outside the software but must be accessible from the inside without forcing the user to shift between the systems. The addition of these features in a software with business intelligence platforms helps users to choose where in business methods, analytics should be embedded.

Architecture of Embedded BI

This would help the users to make out aptful decisions to increase their revenue and incorporate valu to software applications. The elevation in business can't alone be achieved by making out decisions or just by predicting the outcome alone. It is necessary to take actions such as: using the analysis for making out better decisions, efficiently accessing the analysing data and business effectiveness through data integration i.e: integration of analysed data with CRM, ERP, sales, marketing and other such fucntional bodies of the company. ■

About the Authors

Dr. S. Balakrishnan, (CSI Membership 2060000034) is a Professor at Sri Krishna College of Engineering and Technology, Coimbatore, Tamilnadu, India. He has 17 years of experience in teaching, research and administration. He has published over 15 books, 3 Book Chapters, 4 Technical articles in CSI Communications Magazine and over 100 publications in highly cited Journals and Conferences. His professional awards include: Deloitte Innovation Award, Cash Prize Rs. 10,000/-, from Deloittee for Smart India Hackathon 2018, Patent Published Award, Impactful Author of the Year 2017-18, Best Faculty – Computer Science and Engineering, Teaching Excellence Award, I2OR - Bright Researcher Award, Best Outstanding Faculty Award, Best Teacher Award, Best Research Paper Award, Best Book Publication Award and Best Book Chapter Award, Special Contributor Award and Star Performer Award. His research interests are Artificial Intelligence, Cloud Computing and IoT. He has delivered several guest lectures, seminars and chaired a session for various Conferences. He is serving as a Reviewer and Editorial Board Member of many reputed Journals and acted as Session chair and Technical Program Committee member of National conferences and International Conferences at Vietnam, China, America and Bangkok. He has filed/published Patents on IoT Applications. Dr. Balakrishnan is a life member of ISTE, IAENG, IEAE, IARDO, CSI, UACEE, SDIWC and CSTA.

Rahul R. is now pursuing his Undergraduate degree in Sri Krishna College of Engineering and Technology. He has participated in technical events conducted in many colleges and has participated in the international conference on Big Data Analytics.

Big Data Computing as a medium for smooth processing of critical infrastructure in smart cities

Gopal Krishna and Subhash C. Pandit
Netaji Subhas Institute of Technology, Patna, Bihar

Overview

Big data is a notion used for the most tedious and large collection of information in real-time or in batches that traditional modes of computation are not compatible enough to handle them smoothly [3]. Various social networking sites like Facebook, Twitter, LinkedIn produce trillions of data every year. Also, sensor enabled data are generated across various static and mobile sources like flights engines, smart factories, transportation services etc. This exponential growth in data which is highly unstructured is to be first stored and then simultaneously processed to retrieve valuable information that are unavailable to a naked eye. Various challenges are associated with this way of data storage and processing sharing, transfer, visualization, analytics etc. [2] To deal with each of these challenges individually, various open source and licensed tools and frameworks are continuously being introduced.

Big data can be a very powerful tool for personalized, secure and smooth processing of the critical infrastructures available in smart cities like grid, factories, transport, healthcare, water supply and even the internet [1,3]. Big data computing enabled with powerful machine learning tools can help in emergency management and defect monitoring of these critical infrastructures [4]. To ensure effective, real-time and local assessment of the critical equipment, stream processing tools of big data computing prove to be handier and powerful with features of parallelization and data correlation in a secure environment.

The basic structure involved during dealing of these processes is an implementation stack representation with the prescribed sequence of work like data ingestion, data preprocessing, retrieving data intelligence and finally data visualization is shown in Figure 1 below.

Data Ingestion: The structured and unstructured data from heterogeneous sources like sensors, mobile data feeds, and videos are to be ingested using Apache Kafka onto Cassandra DB hosted on cloud environment.

Pre-processing & dimensionality reduction: Fast data is subjected to Cleaning, summarization and finding co-relations of the data for removal of duplicate, incomplete values.

Data intelligence: The fast cleaned data is trained on various machine learning models depending on type of factor involved in the learning. Machine learning model is tested for subsequent classification, clustering and percentages of regression prediction of the instrument reliability and time to live factors. Scikit-learn library of Machine learning and MLlib library of Apache Spark can be simultaneously used to process such critical data involving the critical infrastructure.

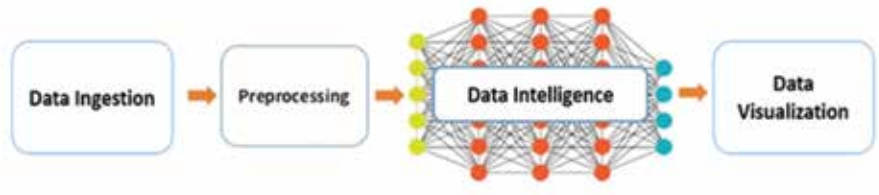


Fig. 1 : Big Data Pipeline for machine extraction from critical infrastructure

Data Visualization: Data visualizations is to be performed using pie-charts, graph plots, histograms to aid the horizon of circumstantial knowledge.

A brief approach for handling the various scenarios involved while dealing with critical infrastructure is described below:

The computation on the various heterogeneous data collected is done in two-folds and responses are generated and also recorded within fixed time-intervals in a day. Firstly, certain kinds of data which are latency sensitive are required to be processed to provide solutions to problems.Secondly, data which can be processed with lesser restriction on time sensitivity will be dealt with separately with better computational environment and enhanced analytics tools. The initial phase of this system is the collection of various heterogeneous data from different data sources like sensors deployed. Most of the data generated may not be require for the current computational architecture but may be stored for extended analysis in the future.The streams of data generated from various sources is ingested using Apache Kafka onto a stable NoSQL storage like Cassandra DB and Mongo DB hosted on cloud. Also, depending on the sensitivity of applications as discussed earlier, certain raw data from specific data source will be sent to the storage system. Then, regression and classification techniques of Apache

Spark MLlib to predict and analyse the state information of devices, networks, sensors or any other type of equipment involved is applied. Finally, a comprehensive view for the large amount of processed big data in a clear and clear and cohesive way is provided Graphical overview of a business model helps in creating a faster observatory analysis and querying at a faster rate. Also, data visualization can be used as a comparative measure for finding the pros and pitfall of this proposed system from the earlier existing ones.

References:

[1] Z. Nyikes and Z. Rajnai, "Big data, as part of the critical infrastructure,"*2015* IEEE 13th International Symposium on Intelligent Systems and Informatics (SISY), Subotica, 2015, pp. 217-222.

[2] Wu R, Huang L, Yu P, Zhou H. EDAWS, "A distributed framework with efficient data analytics workspace towards discriminative services for critical infrastructures". Future Generation Computer Systems. 2018 Apr 30;81:78-93.

[3] Wang T, Bhuiyan MZ, Wang G, Rahman MA, Wu J, Cao J. "Big data reduction for a smart city's critical infrastructural health monitoring". IEEE Communications Magazine. 2018 Mar;56(3):128-33.

[4] Aniello L, Bondavalli A, Ceccarelli A, Ciccotelli C, Cinque M, Frattini F, Guzzo A, Pecchia A, Pugliese A, Querzoni L, Russo S. "Big data in critical infrastructures security monitoring: Challenges and opportunities". arXiv preprint arXiv:1405.0325. 2014 May 1.

■

About the Authors

Mr. Gopal Krishna is currently working as an Assistant Professor in Computer Science & Engineering Department and student branch [F8001991] of CSI Student Branch, Netaji Subhas Institute of Technology, Patna, Bihar. He has 6 years of experience in industry and academia. His area of research are Online Social Network, Machine learning, Data analytics, Wireless Sensor Networks. He has various publications in Machine leaning, Wireless Sensor Networks and Online Social Networks.

Mr. Subhash Chandra Pandit, is currently working as an assistant professor and head in Department of Computer Science & Engineering, Netaji Subhas Institute of Technology, Patna, Bihar. He has 8 years of experience in industry and academia. His area of specialization is online social network analysis, data analytics and intrusion identification and prevention.

CSI CALENDAR 2018-19

Date	Event Details & Contact Information
NOVEMBER 24, 2008	**One Day National Seminar on BIG DATA Analytics & Business Intelligence (BI)** **Venue : Dr. Zakir Hussain Auditorium, Kumharar, Patna-7 Organised by :** IIBM & CSI Patna Chapter Contact : Md. Shams Raza (M) 94308 28918 and Prof. Ganesh Pandey (M) 94314 42560
FEBRUARY 25-28, 2019	**Second International Conference on Advanced Computational and Communication Paradigms (ICACCP-2019)** International Symposium on Computer Vision and Machine Intelligence in Medical Image Analysis (ISCMM) http://symposium.icaccpa.in/ Venue: Sikkim Manipal Institute of Technology **Convener :** Prof. Debanjan Konar, Sikkim Manipal Inst. of Technology, Sikkim, India **Co-Convener :** Prof. Chinmoy Kar, Sikkim Manipal Institute of Technology, Sikkim, India
MARCH 01-02, 2019	**2019 International Conference on Data Science and Communication (IconDSC)** in technical association with IEEE-Bangalore Section, IEEE-ComSoc, Bangalore Section and CSI Division IV, Communication. Submission Deadline: 10 November'18, https://christuniversity.in/icondsc/ Contact : Dr. Samiksha Shukla, 9880462311 samiksha.shukla@christuniversity.in
13-15, 2019	**INDIACom-2019 (IEEE Conference ID: 46181 | SCOPUS Indexed)** 13th INDIACom; 2019 6th IEEE International Conference on "Computing for Sustainable Global Development" Contact : Prof. M. N. Hoda, General Chair, INDIACom-2019, Director, Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM) E-mails: conference@bvicam.ac.in, indiacom2019@gmail.com Tel.: 011-25275055 TeleFax: 011-25255056, Mobile : 09212022066

An Analysis of Data Mining using Python

Ashish
Student, CSE, N.I.E.T, GR. Noida, India
ashishbhardwaj354@gmail.com

Garima Verma
Student, CSE, N.I.E.T, GR. Noida, India
garimaniet1234@gmail.com

Kalpana Dwivedi
Asst. Professor, CSE, N.I.E.T, GR. Noida, India
dwivedi.dolly@gmail.com

Sanjay Kumar Nayak
Asst. Professor, CSE, N.I.E.T, GR. Noida, India
sanjaynoida@rediffmail.com

Data mining study goes for modelling and database Design method. In this research paper we begin by discussing what data mining is, why it developed now and what Challenges it faces, and what types of problems it can address. Those that can master this technology and its methods can derive great benefits and gain a competitive advantage. We will further see in this paper how Python can help in Data mining, various aspects, features and advantages of using Python in data mining in various fields which may improve day to day life.

Keywords- evaluated and interpreted, preliminary programming, PYPL, enormous demand, Intelligibly, unethical.

1. Introduction

The process which takes data (data may be of any kind) as input and outputs knowledge is generally called Data Mining. One of the earliest and the most efficient definitions of the data mining process, which highlights some of its distinctive characteristics, are provided by Fayyad, Piatetsky-Shapiro and Smyth (1996), who define it as "the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data." [1]

As shown in Figure 1 we have random data stored in memory then we select the required data which referred here as target data, after selecting the data comes in pre-processing state. After transformation the data turns into transformed data, now the data is available for Data mining. Data mining turns the transformed data into patterns which can be evaluated and interpreted and use as knowledge. This process is called the Data mining process.

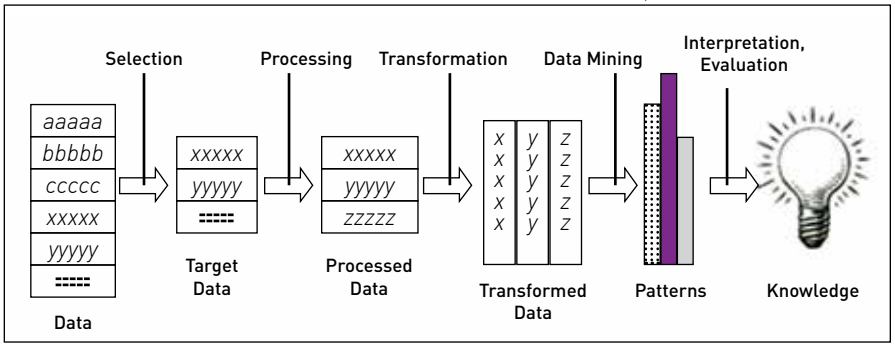


Fig. 1 : The Data Mining Process

2. Using Python for Data Mining

We have noticed a lot of practical reasons for using Python (programming language) in the data mining which are:

1. **Programmers accept Python as a simple language with high readability.** Even non-programmers find it too easy to understand Python in comparison to other programming languages. The language is so clear and simple so that it motivates the user for understand, and the different programming. A Program written in Python (programming language) may also be 2-5 times shorter than other programs written in any other programming languages.
2. **Platform independent.** Python basically runs on the 3 main evaluating platforms which are Microsoft windows, Apple Mac, Linux as well as on a wide variety of other platforms. Python is very different programming language but can be useful tools for modern developers. If you want to learn programming, you shall opt Python because of its reliability factor.
3. **Interactive Program.** This is basically the type of programming in you don't require to introduce much details as it is easier to programming is when you compile the code and run evaluates it and things are gone. Intelligibly, this kind of programming is much better, because it makes programming easier, looking at data easier and trying out code on it [2].
4. **General Purpose Language.** Python is a kind of general purpose language which can be easily used to a huge variety of tasks other than data mining, e.g., applications (users), computer system administration, gaming, web development, psychological presentations based on experiments, recording and many more.
5. **Large community.** Python

(programming language) has a tremendous community and has become more popular in less time. Some of the indicators around the world have testified to this [3]. The Popularity of Language Index (PYPL) sets its ranking as programming language's community, On Google search for ranking, volume provided by Google Trends puts Python in the third rank after PHP and Java. As per to PYPL the dignity of Python has grown continuously since 2004. TIOBE constructs another rank producer that puts Python in 6th rank. This rank produces on basis of the number of engineers worldwide, third party courses. And from various surveys it was found that Python is the rising programming language in terms of Git-Hub Projects Stack Overflow result. Moreover, in a survey in 02014 Python was ranked the easier and popular programming language among all at United States universities for teaching preliminary programming [4].

6. **Quality.** A company named Coverity has found that Python code errors in between its 400,000 lines of code, but that the error rate is very fewer in comparison to other projects that are open source. [4]

3. Aspects of Data Mining

1. Uncertainty Handling
2. Dealing with Missing values
3. Dealing with Noisy data
4. Efficiency of algorithms
5. Discovering Useful or Interesting Knowledge
6. Incorporating Domain Knowledge
7. Size and Complexity of Data
8. Data Selection
9. Understandability of Discovered Knowledge
10. Consistency between Data and Discovered Knowledge.

4. Comparison between different programming languages used for Data Mining

Here you can see the comparison of different languages that are frequently used for Data Mining purposes. Here we take a comparison table from the most recent survey. This comparison helps to understand this concept and need more precisely.

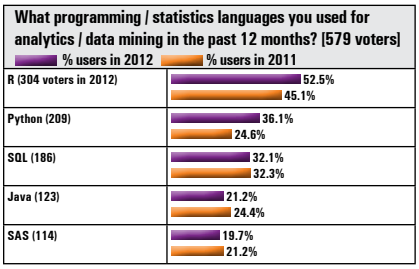


Fig. 2 : Comparison table

As it is clear from above survey, the maximum increment rate is of Python. This survey is proof that python will grow at the same rate in future. Hence, we must focus on python (the fastest growing language) for doing Data Mining and for other programming works.

5. Limitations and Disadvantages

5.1.1 Privacy

Peoples bother about their personal privacy that has been increasing vastly especially on increasing influence of internet on social networks, e-commerce, forums, blogs, etc. Because of the issues regarding privacy, peoples are afraid that personal information of them is being collected and misused, which causes a lot of problem. In order to understand, the purchasing behaviour trends about their customer; the information is being collected. However, we all know, business is like getting mingle in a fake relationship. It might be hold by any other person or vanishes one day. Nowadays, the information is being taken and leaked out in media.

5.1.2 Security

Security is a major issue. Businesses takes information about employees working with the company and their customers that comprises of social status, security detail, life, habits sometimes etc. But the most important questions arise, whether the information arises, whether the information is taken by the companies is kept in care or not? There are much well-known incidents where, data of customers are being stolen by the hackers such as Microsoft Company, Ferrari etc. That generally includes loads of financial and personal information, the identity and thefts cards are the common issues.

5.1.3 Misuse of inaccurate information

The information collected via Data mining for ethical purpose can be misused and might be sold or leaked to unethical people or business for their mean advantages to fulfil their own mere interests.

In addition, The Techniques of Data Mining are not perfectly accurate. Therefore, conclusion is that if improper information is used for decision-making, resulting in some serious major problems in future. [5]

6. Conclusion

Data Mining's use in enrolment management is an equitably new kind of development in this field. Current Data Mining particularly done on simple categorical and numeric data as per requirement. In future, Data Mining will include data types that will more complex as they are increasing day by day. In addition, for any type of replica that will designed, further changes or improvements that may can done by testing other factors, variables and relationships that affects Data Mining. Research in Data Mining will lead in having new methods in future to regulate the most fascinating characteristics in the data. As replicas are developed and tested, they can be further used as equipment in enrolment management.

7. Scope

Data Mining's main motive is detecting useful patterns from database or we can also call it "extracting useful information" from data (database) which is whole do not look as useful or perceptive. The applications of Data Mining are in every field exist today like from e-commerce to Oil exploration, from space agencies like NASA, ISRO, ROSCOSMOS, to social networks like Facebook, twitter, WhatsApp, Instagram etc. There is an enormous demand of data-miners alias data scientists. So, a course in the field of Data Mining worth doing, along with course practical implementation of concepts is equally important to understand what really can be done with the help of Data Mining. Seeing the demand of data scientists, every second company today is looking for data scientists, so we think job

won't be an issue in future. There will be a huge demand of Data scientists in upcoming future, so the peoples who are going or thinking to go in this field should not have to take about their future or livelihood. You will have a great future ahead and secured also.

References

[1] Handbook of Technology Management, H. Bidgoli (Ed.), John Wiley and Sons, 2010.

[2] https://yosefk.com/blog/code-data-and-interactive-programming.html.

[3] http://www.tiobe.com/index.php/content/paperinfo/tpci/index.html.

[4] Reference Google facts.

[5] http://www.zentut.com/data-mining/advantages-and-disadvantages-of-data-mining

■

About the Authors

Mr. Ashish is a student of Computer Science Engineering from NIET, Gr.Noida and he is writing his first research paper in the area of Data Mining using Python.

Education and Credentials: He is pursuing his Bachelor of Technology (CSE) from Noida Institute of Engineering & Technology. He had completed his senior secondary education from Rakhee Public School.

Kalpana Dwivedi is an Asst. Professor of Computer Science Engineering in NIET, Gr. Noida. She has 6 years of Experience in Teaching and research work.

Area of Interest: Big Data Technology like Hadoop, Splunk, Mongo DB, Cassandra

Education and Credentials: Her academic qualification is M. Tech (CSE), B. Tech (CSE). Her research paper is published in IEEE.

Role of ICT in Education

S. N. Kakarwal
Professor in Department of Computer Science & Engineering of P.E.S. College of Engineering, Aurangabad

R. R. Deshmukh
Professor in Dept. of CSIT, Dr. B.A.M. University, Aurangabad Maharashtra

D. V. Kurmude
Associate Professor and Head in Department of Physics at Milind College of Science, Aurangabad, Maharashtra

What is ICT?

ICT (Information Communication Technologies) is the use of communication and computer based technologies that enables the user to store, retrieve, transmit, receive and manipulate digital information. From last two-three decades the use of Information Communication Technologies has greatly impacted in various sectors such as education, business, governance etc. And it has changed many aspects of the lives. Today's world is computer world and this world is rapidly moving towards digital information. So, the importance of ICT in education is increasing. This article gives highlights various impacts of ICT in education and future of ICT based education. The article also argues the role of ICT in transforming teacher-centered learning to Student centered learning.

Introduction

As education is need of society because the education plays the vital role in building the society. The quality of education determines standard of society. The educational quality helps to empower the society in all aspects by bringing new thoughts, new technologies and so many such things. There are number of effective teaching and learning methodologies that are in practice today. But these traditional methodologies are teacher centric. In traditional approach, we have particular course structure / syllabus, subject wise textbooks & reference books and we follow that for years. Here teachers teach in lectures either using board or presentation. Teachers used their lesson plans, tutorials etc and the assessment of student's performance is done using assignment, exams etc.

Traditional teaching-learning process

In conventional teacher-centric teaching-learning process, students just study the contents of syllabus that is not sufficient for bringing and implementing new ideas and new technologies. In such system students only learns what their teacher teach in lecture. They cannot choose topic of their choice whenever they want. Students have to attend the lectures even though they are bored or not interested in particular topic. To survive in this dynamically changing technological world, student-centered learning process is necessary where student will have anytime access to required information by the use of web as a source. Students can learn anytime and at any place. Students can select the experts from whom they will learn etc. Here comes the role of ICT in the education sector! Learning Process will be problem based learning.

Need for Competency based learning

To compete with this rapidly growing world we need to improve quality and structure of the education system by enforcing competency and performance based approach by including advance technology and practical based approach in learning process. With such Student-centered leaning process, student has access to any type of information. In Such system student will learn variety of things that are beyond the syllabus with Real life examples. In such system Teachers are not content experts but they acts as a mentor/guide to the students.

ICT as change agent

ICT plays a vital role in standardizing the quality of education. Use of ICT helps to transform teacher centered learning to competency based learning. As ICT supports independent learning, Students get deeply involved in a particular activity. To bring this changes Educational board and universities should also choose ICTs as a learning technology. The following are the advantages of ICT based learning over conventional learning process:

a) **Students Centered Learning:** With the help of web as source students can learn anything without the need of teacher. Students can choose the mentors from whom they want to learn. In Such system teachers are not content expert of the particular subject but rather they are mentor for the student.

b) **Supporting Knowledge Construction:** In The conventional teaching-learning process teachers plan the things and then delivered those things through lecture. But this system does not take into account individual Student's interest, their understanding level etc. This interests and understanding level can be taken into account with the help of ICT based learning. Use of ICTs supports knowledge construction and also facilitate many opportunities through their provision and support for resource based, student centered learning.

c) **Any place learning:** The use of ICT allows students to learn at a distance. In conventional system students have to come and attend lectures in classroom. But now with the help of ICT Student can learn from a distance and according to their convenient time. This type of learning not only provides convenience but also saves the cost associated with travel and time away from work. With these modern technologies learner gets the opportunities to enroll the courses offered by external institutions rather than those

situated locally. Such freedoms of selecting course of our choice also support the delivery of programs with units and courses from a variety of institutions. Students can learn single topic through lecture series of number of different institutions.

d) **Any time learning:** ICT based learning system also remove the temporal constraints by providing opportunities for students to undertake education anywhere, anytime and at any place. Through online technologies Students can learn in their schedules and slots. Learners can participate in learning activities according to their time and such freedoms provide the opportunities for students to participate in formal programs. Unlike conventional learning where teacher teaches in classroom only once and if students want to listen the lecture again and again, then it is not feasible. But with today's Mobile technologies and communications technologies students can listen the lecture any number of times. Teaching and learning is now possible for 24x7. For example From April 2001 MIT has kept all there learning materials (including PPT's, Notes, Videos lectures, question paper etc.) freely available online for learning and sharing purpose.

Role of ICT in enhancing the development of literacy

ICT is not only limited to computers and communication technology but ICT is the broad term which consist of Radio, TV, Telephone, Mobiles, Fax, Satellite, CD-ROM, Internet etc. Learning based on new ICT technology is also referred as e-Learning. With the help of e-learning learner can access any type of information related to its work such as statistical information of stock market or farmer can see weather forecasting to improve their agriculture production by using communication devices.

Merits of e-learning:

1) We can access e-learning programs from anywhere and at any time due to wide network.
2) We can share information with other peoples by using Cloud computing services.
3) Cloud plays very important role to store and exchange the information with help of privacy preservation techniques.
4) Learner can learn anything without teacher from anywhere i.e. distance learning mechanism.
5) Reduces adult illiteracy rate, with sufficient emphasis on female literacy.

Role of ICT in Higher Studies

Now a days ICT is playing very significant role in higher studies or in research also such as Technical institutes are using ICT to improve the skill of students in various areas. Student can access lecture notes or study material like Video lectures and e-Books.

Objectives of ICT in Higher studies:

1) To improve the quality of education and soft skills of students
2) To provide e-resources to students easily and in low costs.
3) To promote technology awareness.
4) To reduce information sharing time between teacher and students.

Role of ICT in School and non-formal education

ICT provides all time interactive learning with web based information sources and promote change in teaching learning method. It also reduces the costs requires for textbooks. Now a days in school ICT provides cost-effective learning method with the falling cost of hardware, internet access, maintenance etc. School students can learn very easily in cost effective manner. With the help of ICT technologies it becomes possible to Support non-formal education for out of school children and adults.

Infrastructure required to setting up ICT in schools or Colleges

Some hardware and software required to set up ICT in institutes are given below

1) Computer Labs
2) Network devices having sufficient bandwidth
3) Peer to peer and Client server network
4) High speed Wired or wireless internet access
5) Operating system softwares
6) Application softwares such as word press, spreadsheets, Graphics softwares
7) Web based softwares such as Web browser, java in build browsers
8) Projectors, LCD display devices

■

About the Authors

Dr. S. N. Kakarwal [CSI-F8000602] is presently working as Professor in Department of Computer Science & Engineering of P.E.S. College of Engg., Aurangabad, MS-India. Her research interests include Image Processing, Pattern Recognition. She has published 13 papers in International Journals, 25 papers in International Conference and 11 papers in National Conferences. She has published four articles in CSI magazine. She can be reached at s_kakarwal@yahoo.com

Dr. R. R. Deshmukh, [CSI- 00100518], is currently working as Professor in Dept. of CSIT, Dr. B.A.M. University, Aurangabad, (MS), India. He has been elected as sectional member of ICT section of Indian Science congress Association. His areas of specialization are Human Computer Interaction, Digital Speech Signal processing, Computational Auditory Scene Analysis (CASA), Neural Networks etc. He can be reached at ratnadeep_deshmukh@yahoo.co.in

Dr. D. V. Kurmude is [CSI-2010001037] presently working as Associate Professor and Head in Department of Physics at Milind College of Science, Aurangabad. MS-India. He has published more than 20 research papers in International Journals and conferences. Authored one book published by international publisher.

Computer Society of India™

Education Directorate, National Headquarters,
CIT Campus, IV Cross Road, Taramani, Chennai – 600 113.
Phone: 044-2254 1102 / 03 / 2874 **Website:** http://www.csi-india.org
Email: ed@csi-india.org **Copy to:** csipromotions@csi-india.org

Form – I
Version 2.4
w.e.f. 01st July. 2017

Application Form for Institutional Membership (Academic and Non-Academic)

Complete this application form and send it along with the payment by Speed Post / Courier to CSI, Chennai (above address)

(Put ✓ in appropriate box)

Academic (Colleges / Universities)	**Non-Academic** (Corporates / PSU / Govt.)

School	**Polytechnic**	**College**

New	**Renewal**

1) If an existing member, proposing to renew, enter your Institutional Membership No. _____
2) Period of Membership applied for: 01 / 02 / 03 / 04 / 05 / 10 / 15 / 20 Years **(Tick one, which is applicable)**
3) Name of the Institution (BLOCK LETTERS) _____
   _____
4) Postal Address (BLOCK LETTERS) for communication: _____
   _____
   City _____ State _____ Country _____ Pin Code _____
5) Phone with STD Code _____ E-Mail ID _____
6) Details of the HEAD of the institution (Nominating Authority):
   Name: Prof. / Dr. / Mr. / Ms. _____
   Designation _____ E-Mail ID _____ Mobile No.: _____
7) **Nominees' details – Faculty / Staff Members:** (Max. 03 Nominees for Academic and 04 for Non-Academic): They will be offered free Professional Membership of CSI to be valid for the period of the Institutional membership with all the rights including voting. While nominating, please nominate such a person, who is currently not the member of CSI, to enable him to become a part of the organisation. All the nominee members are also required to fill up their Form No. II, applicable for Individual membership, so **that their detailed information can be recorded by CSI,** and send the same along with this form.

Nominee No.	**Name** (Prof. / Dr. / Mr. / Ms.)	**Mobile No.**	**E-Mail ID**	**CSI Membership No.** (to be allotted by CSI office)
01				
02				
03				
04				

Note: Additional paid nominees can be added by adding the additional payment to be calculated as applicable to the individual members.

8) **Payment Details:**

Mode of Payment			
If the payment is to be done by offline DD / Cheque, issue it in favour of "CSI Education Directorate" payable at Chennai.		**If the payment is to be done online by Net Banking / RTGS / NEFT / etc., then following details may be used for payment and transaction detailed be filled in.**	
Cheque / DD No.		Transaction ID	
Amount (Rs.)		Amount (Rs.)	
Date		Date	
Drawee Bank		SB A/C. No	1 0 7 9 2 4 6 3 7 4 5
Branch Name & City		Branch Name	SME Branch, Adyar, Chennai – 600 020
		Bank Name	State Bank of India (SBI)
		IFSC	S B I N 0 0 1 3 3 6 1
		Please send the photocopy (or) scanned copy of the Net Banking Transaction Slip at ed@csi-india.org; csipromotions@csiindia.org	

9) Mention name of Region & Chapter nearest to you: Region _____ Chapter Name_____

10) **Code of ethics - Undertaking:**

We affirm that as a CSI member, we shall abide by the code of Ethics of the Computer Society of India (CSI). We, further, undertake that we shall uphold the fair name of the Computer Society of India by maintaining high standards of integrity and professionalism. We were not a member of CSI earlier / We were a member (Membership No. _____) earlier and membership ceased without prejudice. We are aware that breach of the code of ethics may lead to disciplinary action against us under the Constitution, Byelaws and Rules framed from time to time. We hereby confirm that we shall be bound by any decision taken by the CSI in all matters. I hereby convey my consent to receive the CSI publications in soft copy form and any other information about the activities of the society by email or by SMS on my Mobile number, from time to time, by the society or the members of the society.

Date: / / Signature and Name of the Head / the Nominating Authority

Place: (Seal of the Institution)

FOR OFFICE USE ONLY

Application received date : _____ Received By : _____

Application processed by : _____

Membership No. | | | | | | | | |

(Membership subscription fees details for the information of the applicant , not to be attached along with the Application Form to be sent to CSI)

Institutional Membership Subscription Fee

(Academic and Non-Academic) w.e.f. 01.07.2017

(The membership Period is on Rolling Year basis)

Details of the Membership Categories	Period – Wise Membership Fee + GST, as applicable (in ₹)							
	01 Year	02 Year	03 Year	04 Year	05 Year	10 Year	15 Year	20 Year
Institutional Members (Academic) With 03 free Nominees	6,000	11,000	16,000	21,000	25,000	48,000	70,000	90,000
GST @ 18%	1,080	1,980	2,880	3,780	4,500	8,640	12,600	16,200
Total Membership Fee	**7,080.00**	**12,980.00**	**18,880.00**	**24,780.00**	**29,500.00**	**56,640.00**	**82,600.00**	**1,06,200.00**
Institutional Members (Non-Academic) With 04 free Nominees	10,000	19,000	28,000	36,000	45,000	85,000	1,25,000	1,50,000
GST @ 18%	1,800	3,420	5,040	6,480	8,100	15,300	22,500	27,000
Total Membership Fee	**11,800.00**	**22,420.00**	**33,040.00**	**42,480.00**	**53,100.00**	**1,00,300.00**	**1,47,500.00**	**1,77,000.00**

Individual and Life Membership Subscription Fees

(The membership Period is on Rolling Year basis)

1) **Individual Membership Fee:**

Membership Category	One Year	Two Years	Three Years	Four Years
Within India	₹ 1000 + 18% GST = ₹ 1,180.00	₹ 1800 + 18% GST = ₹ 2,124.00	₹ 2600 + 18% GST = ₹ 3,068.00	₹ 3400 + 18% GST = ₹ 4,012.00
Outside India	$60	$ 110	$ 150	$ 180

One Day National Seminar on Cyber Security & Cyber Forensics

Organised by Department of CSE & CSI Student Branch

Netaji Subhas Institute of Technology, Bihta, Patna

October 6, 2018, Department of CSE and CSI Student Branch, NSIT, Bihta, Patna has organized One Day National Seminar on Cyber Security & Cyber Forensics at its campus. Prof. (Dr.) A.K. Nayak, Vice President of CSI & Chief Guest inaugurated the seminar by lighting the lamp along with other dignitaries present. The guests were welcomed by Mr. M. M. Singh, Founder Secretary, NSIT. Prof. M.P.Tripathi, Chariman, BOG, NSIT has introduced about the institute. Mr. Krishna Murari, Registrar, addressed the gathering with best wishes for successful deliberation. Dr. Sujit Singh, Dean Academics, NSIT addressed about the academic activities. Mr. Gopal Krishna, SBC, CSI student branch introduced the seminar. Ms. Puja Bharti and Ms .Bipra, students of CSE-4th year were the anchors of the inaugural session.

Prof. M.N. Hoda, director, BVICAM addressed as the key-note speaker. Mr. Shailesh Shrivastava delivered his talk on secured e-transaction. Dr. Prabhat Kumar, CSI State Student Coordinator of Bihar and HOD, CSE, NIT Patna presented his talk on security in IoT. Dr. J. P. Singh, Assistant Professor, NIT Patna delivered his talk on Cyber Forensics. Topic of Dr. S.C. Yadav, Associate Professor, Central University of Jharkhand was Cyber space, Cyber Crime and Cyber Security. Md. Shams Raza Incharge, XCCS spoke on Estimation and Optimization of Cyber Security for Business and Trade to support safe financial transaction.

More than 500 faculties, industry persons, research scholars and students from more than 30 institutes and organizations from all over India has participated in the seminar. Some of the notable among them were IIT Patna, NIT Patna, BIT Mesra, Central University of Jharkhand, Patna Women's College, Magadh Mahila College, LNMI Patna, TMBU Bhagalpur, IGNOU, Bakhtiyarpur College of Engineering, B.S. College, Marwari College etc. Various activities were conducted in the seminar like experts talk, live hacking session as industry presentation by ICETL, paper presentation and poster presentation. Winners of competitions were awarded by certificates, medals and cash money prizes. This seminar was also supported by ICETL and Zebronics.

Mr. Subhas C. Pandit, HOD, CSE, NSIT, Patna proposed the vote of thanks. All faculties of NSIT actively participated to make this function a grand success. Few notable among them were Mr. Aditya Shekhar, TPO, Mrs. Shradhha Pandit, Mrs. Sarita Chaudhary, Mr. Manish Kumar, Mr. Rajesh Kumar Mishra, Mr. Ramakant Singh, Mr. Pradeep Kumar, Md. Mohtab Alam, Mr. Triloki Nath, Ms. Preyashi Singh, Mr. Hareram Ray, Mr. Arun Kumar Singh, Mr. Ranjeet Kumar, Mr. Rajni Ranjan, Dr. J. Dalai, Mr. R.K. Chaudhary, Mr. Ashish Pathak, Dr. Reema Dhar, Mr. Kundan Singh, Md. Hasmat Ali, Ms. Surbhi Priya, Ms. Neha Chauhan, Ms. Pallavi Singh, Rakesh Kumar, Mr. Prakash Kumar, Mr. Brijesh Kumar, Ms. Prakashika Kumari.

Prof. (Dr.) A. K. Nayak, VP, CSI while delivering the inaugural address

Mr. M. M. Singh while delivering the welcome address

Faculties and Students of CSE, NSIT

National Anthem

Participants in Auditorium

Report on Workshop: "Cyber Security & Data Privacy"

Dr. R. C. Tripathi
Professor School of CSE & Chairman, CSI, Noida Chapter

Prof. Neetu Singh
Workshop Convener HBS

Date : **13th, October, 2018**
Venue : **Hierank Campus: Auditorium**
Name of the Event: **Workshop on Cyber Security and Data Privacy**
Number of participants in the event: **500**

CSI Noida Chapter & IT Department, HIERANK Buisness School, Sector 62, Noida, UP had jointly organized a WORKSHOP on the topic: Cyber Security And Data Privacy in the campus of HIERANK Business School on 13th October 2018 in which more than 500 participants from the Colleges of Delhi NCR participated.

The event started with Lightning of Lamp by dignitaries' followed by welcoming and introductory speech by Prof. (Dr.) Rajesh Sahay (Chairman Hierank Business School).

Fig. 1 : Prof. (Dr.) Rajesh Sahay (Chairman Hierank Business School) addressing the workshop

The occasion was graced by the presence of Chief Guest Prof. (Dr.) S. K. Kak Former VC, CCS University Meerut UP & AKTU Lucknow UP.

Fig. 2 : Prof. (Dr.) S K Kak Addressing the workshop

Mr. Anuj Aggarwal (Founder Chairman CSI Noida Chapter), Mr. Dinesh Bareja (Cyber Expert), & Dr. R. C. Tripathi (Professor, School of CSE, Lingayas University & Chairman CSI Noida Chapter), Mr. Pulin Kumar (Director, Adidas), Dr. Parmanand (Dean, SET Sharda University), Dr. Nitin Rakesh (H.O.D., CSE Sharda University), Ms. Suhasini Gotambre (IA & AS), Dr. Rajendra Singh (Professor, Monad University), Dr. Manish Saraswat (H.O.D., M.C.A., Geetanjali Institute Udaipur Rajasthan), Mr. Rahul Kumar (Delhi Police) and Mr. Dinesh Batra (Sr. Vice President Hindustan Power), Dr. Amit Srivastava (Director HBS), Dr.Pragati Saxena (Dean HBS) & Ms. Neetu Singh (Convenor HBS) along with IT students and faculty members of the various colleges of Delhi NCR.

Fig. 3 : Prof. (Dr.) Rajesh Sahay, Prof (Dr.) RC Tripathi, Prof. (Dr.) S K Kak, Prof. Parmanad & Prof. Dr. Amit K Srivastav during the workshop

Chief Guest Prof. (Dr.) S. K. Kak Former VC, CCS University briefed the audience on the security issues regarding the AADHAR cards. Explaining the vulnerabilities we face by the hands of hackers also providing with measures of prevention to secure out Personal Data. Mr. Anuj Aggarwal (Founder Chairman CSI Noida Chapter) & Dr. R. C.Tripathi (Professor, School of CSE Lingays University & Chairman CSI Noida Chapter) enlightened everyone about Cyber Security and Data Privacy issues with detailed power point presentations. Highlighting the risk of personal data steeling and leakage over various social networking sites. Ms. Suhasini Gotambre(IA & AS) and Mr. Pulin Kumar(Sr. Director Legal & Corporate Affairs, Adidas India Pvt. Ltd.)also briefed the audience about the measures that can applied to prevent unauthorized access to computers, databases and websites protects data from corruption.

Fig. 4 : Prof SK Kak, Mr Anuj Aggarwa & Chairman HBS in Saraswati Puja

The Second half of the session included Hands- on Training by Mr.Devanshu Shukla (Director Hackveda.in) from VMDD Technologies, Rohini New Delhi & Mr. Dinesh Batra informing students about the pros and cons of hacking and guiding them with the basics. The session was concluded by honouring the guests with a token of appreciation by the Prof. Rajesh Sahay (Chairman Hierank Business School).

Fig. 5 : Prof Dr. R C Tripathi addressing the workshop

The Vote of thanks was conveyed by Prof. Neetu Singh (Convener of the Workshop) The event was convened by Prof. Neetu Singh and co-convened by Ms. Ruchi Srivastava along with student committee compromising of Mr.Abhijeet Sharma, Ms. Vishakha Tripathi, Ms. Sakshi Verma, Mr. Nitin Singh Rawat, Mr. Suraj Mishra, Mr. Nikhil Kumar Varshney, Mr. Suman Chakraborty, Ms. Shweta Maurya, Mr. Shivanshu, Mr. Ritik Jha and Ms. Pooja Singh. The event was highly appreciated by all the attendees. The Event ended with Valedictory Session.

Computer Quiz (CQ - 2018)

(Computer ville)

R K Vishwakarma
Chairman, CSI Gwalior Chapter, Gwalior (MP), Defence R & D Establishment, Jhansi Road, Gwalior-474002

The Computer society of India, Gwalior Chapter and Department of computer science and engineering, IPS Collage of the Technology and Management, Gwalior has organized a computer quiz for the students of High school & Higher secondary schools of the Gwalior region. This quiz was a part of an event of IPS group of colleges - "Techno Parv 2018". This time this CQ-2018 was held in two phases. The first phase was the offline round and it begins from 14th Sept. 2018. In this off line round, objective based written test was held at various schools. Top three students of the respective schools were selected and they represent their school as Team. After selection of teams from different schools a online and final round was conducted at IPS Group of colleges, Gwalior on dt. 01.10.2018. Thus, more than 500 students from the various schools of Gwalior and outside (Shivpuri) were participated in this quiz. The Chairman Er. R.K. Vishwakarma, Immediate past chairman Mr. J.S. Bhide, Secretary Mr. Dilip Hayaran, Treasurer Mr. Ravindra Gadkari, Director of IPS college Dr. Arun Tygi, Principle Dr. P.S. Chauhan and CEO of the IPS Group of colleges, Mr. P.K. Gosh along with coordinators of the various schools were present during this occasion. Inaugural speech has delivered by HOD Computer Science, Dr. Pankaj Goyal. The anchoring of this quiz was beautifully & successfully conducted by faculties duo Ms. Neha and Ms. Prgya. After completion of all the rounds of this quiz the overall winner was the No.1 Air Force School, Gwalior. Winner and runner teams were awarded by a trophy, certificate and a cash prize. Organizing team of quiz divided the schools according to their zone. So, Winner of zonal category also awarded by cash prize and a certificate. The participation certificate given to each and every team of the quiz.. Chairman, CSI Gwalior Chapter, Er. R.K. Vishwakarma has delivered the concluding remark for this whole event and vote of thanks presented by Prof. Kanunendra Verma.

National Cyber Security Summit 2018 – A Report

ON 6th October 2018 at TCS WANDERRS CSI Kolkata Chapter organized National CYBER SECURITY Summit 2018.

The theme Cyber Security based on the perspective of Cloud, Blockchain, IoT & Big Data and AI. Awareness of new threats and how to mitigate them. The goal was full length discussion on Cyber Law and its outreach.

Honorable Chief Guest Mr. Debashis Sen, IAS, Additional Chief Secretary, Dept. of IT, Govt. of WB & Chairman, HIDCO, Honorable Mr. R. N. lahiri, convenor of this Summit, Dr. Amlan Chakrabarti, Dean, Faculty of Engg. & Tech., Professor & Director, A. K. Choudhury School of IT, University of Calcutta, Mr. Suresh Menon, VP & GM (India East), TCS, Mr. D P Sinha, CSI Fellow & Summit Convenor inaugurated this Summit . Mr. Gautam Hajra, Organizing Chair and Mr. Sourav Chakraborty, Secretary of CSI Kolkata Chapter were on the dias.

Mr. R. N. lahiri giving welcome address about cyber security in different aspects.

Mr. Debashis Sen, IAS, Additional Chief Secretary, Dept. of IT, Govt. of WB highlighted the role of Cyber Security in Current purview.

Dr. Amlan Chakrabarti with the help of case study explained what are the best practices to follow in organizations and personal life to keep ourselves safe.

Mr. Suresh Menon gave some real life case studies of India and abroad.

Mr. D P Sinha pointed out the need behind Cyber Security as well as this type of Summit.

Keynote Address was presented by Mr. Subhendu Das, Scientist-F, Head, eSecurity Lab, ERTL, STQC IT Services, Ministry of Electronics & Information Technology, Govt.of India.

In the second session, Prof. Chandan Majumdar, Professor, Dept. of CSE, Jadavpur University nicely highlighted on Interplay of Information Security and Privacy at the Enterprise Level. Then,

Mr. Koushik Nath, Lead, Cyber Security (IN & APJ), CISCO briefly touched upon the subject" Making Businesses Secure in a Digital World".

Mr. Abhijit Biswas, Sr. Technical Consultant, Trend Micro nicely explained his topic: Dealing with Changing Threat Landscape.

In the second half Mr. Sitaram Chamarty, Head, Security Innovation Lab, explained what is Security and Privacy.

Mr. Susobhan Mukherjee, Chairman, Infosec Foundation and SIG Head, NASSCOM nicely said on his Topic: Managed Defense and Security Orchestration.

last speaker was Dr. Amit Chaudhuri, Associate Director and Group Head – ICT & Services, C-DAC Topic. He shared his experience on Information Security Awareness in Present Indian Context & International Perspectives.

The last event was Panel Discussion. The topic was "Cyber Threats - How secure we are in current IT scenario". In this event, four renowned persons of different fields participated. They were Mr. Bibhas Chatterjee, Public Prosecutor, Govt. of West Bengal, Mr. Kuntal Siddharth, Kolkata Police Cyber Crime, Mr. Ramesh Dube, CISO, UCO Bank and Mr. A Parida, DGM I/c (C&IT), SAIL. Mr. Bibhas Chatterjee, Public Prosecutor, Govt. of West Bengal, Mr. Kuntal Siddharth, Kolkata Police Cyber Crime discussed different types of real life Cyber crime and how to combat it. Mr. Ramesh Dube highlighted the different aspects of Cyber security for about one hour. Mr. A Parida briefly explained the different types of security measures taken by SAIL.

Finally, Vote of thanks given by Mr. Gautam Hajra, Vice Chairman of CSI ,Kolkata Chapter. In his speech , Mr. Hajra said that an event like this cannot happen overnight. He mentioned that the wheels started rolling weeks ago. He thanked all the advertisers who helped CSI, especially TCS. He also mentioned the name of team members who relentlessly toiled to make this program a grand success. He said that this type of Summit requires planning and a birds eye for details. He Gave mementos to the following CSI members like Snehashish Bannerjee, Aniruddha Nag, Mr. Sourav Chakraborty, Siddhartha Roychowdhury, Dr. Madhumita Sengupta and Sekhar Sarkar for their outstanding contribution in this Summit.

Almost 100 students and 60 Industry delegates were attended this Summit.

Computer Society of India™

Call for Proposals from CSI Student Branches to organize

National / Regional / State Level CSI Student Conventions during the year 2018-19

Computer Society of India (CSI) organizes National, Regional, and State Level Student Conventions annually, to enhance the awareness on technological developments and applications, and foster creative professional orientations among the academic community. The Conventions, held at Student Branches, offer excellent opportunities to the students to manifest their technical proficiency and prowess through paper presentations, discussions and extensive interactions with peers and pioneers.

CSI invites Proposals from Student Branches to conduct the National/Regional/State Level Student Conventions to be held during the academic year 2018-19 (April to March).

Criteria: The proposing Student Branch should be very active, with a track record of several CSI activities, and be in good standing through the years 2017-18 and 2018-19.

The proposals for convention will be evaluated, broadly based on the parameters given below:

a) Number of years of continuous valid Student Branch at the college (without break)

b) Average student strength over the past three years

c) Number, quality and level of activities at the student branch

d) Prompt submission of activity reports and financial accounts

e) Ability to attract good speakers from Industry

f) Availability of infrastructure and other resources

g) Financial strength and potential

h) Accessibility and other general conditions

Schedule:

State Student Conventions: To be conducted on or before 31st January 2019

Regional Student Conventions: To be conducted on or before 28th February 2019

National Student Convention: To be conducted only on 6th March 2019 (CSI Foundation Day)

All the National, Regional & State Student Conventions are to be completed according to the above schedule.

The CSI Student Convention Manual describes the guidelines and norms to conduct the student conventions.

The Proposal: Interested Student Branches are requested to send electronic proposals in the prescribed format with all necessary data, including the information stated below.

a) Type of convention proposed: National/Regional/ State level

b) Proposed dates – please indicate two sets of dates

c) A statement of case why the SB should be considered favourably for the proposed event

d) Signed undertaking by the Head of the Institution to provide all the required support (Document with scanned signature)

e) Name & contact details of SBC and the coordinator-designate for the proposed convention

How to send: The Student Branches may send the proposals in the prescribed format on or before 23rd November 2018 to CSI Education Directorate at admn.officer@csi-india.org/ mgsekaran1962@gmail.com) copy to Vice President & Chairman Conference Committee at aknayak@iibm.in

Selection Procedure: CSI Education Directorate scrutinize the applications and shall be sent to the respective RVPs for recommendations. Final selection will be made by the OBs under the Chairmanship of the Vice President & Chairman Conference Committee.

CSI Support: CSI extends partial financial assistance, in accordance with the availability of budgetary resources. CSI also supports the publicity efforts for the Conventions.

Convention Helpline: CSI Education Directorate shall be pleased to offer any information or help on the convention. Please do contact Mr. Gnanasekaran at 98403 41902 for any assistance.

Prof. A K Nayak
Vice President & Chairman Conference Committee

Dr. Santosh Kumar Yadav
Hon Secretary

Computer Society of India™

Application for hosting
National / Regional / State Level Student Convention

Full Name of the Institution						State : Region :
Complete Postal Address						
Institution Phone Number						
CSI Institution Membership No. & Commencement date						
Contact details of the Head of the Institution/Principal	Name					
	Mobile		Email :			
SBC Details	Name					
	Mobile		Email :			
Type of Student Convention (tick the appropriate column)	National	☐	Regional	☐	State	☐
Proposed Date(s)						
Alternative Date(s)						
Theme Name (3 alternatives desired)	1. 2. 3.					
Proposed Events						
Estimated Revenue						
Estimated Expenses						
Estimated Audience Size						
Event Chair & Contact Details						
Program Chair & Contact Details						
Organizing Chair & Contact Details						
Any other information						
Recommended by the Head of the Institution / Principal						
Signature & Seal:						

Computer Society of India

Call For Nominations
for CSI National Elections 2019-2020/2021

Dear CSI Members,

Under CSI Byelaws (Section 5: Nominations and Elections) of the Computer Society of India, the Nominations Committee (NC) is required to invite appropriate groups of Members to submit names of voting members for considering them for the various elective offices of the EXECUTIVE Committee & Nominations Committee (NC)

Members are accordingly invited to submit names of Candidates who are valid voting members of high professional standing, Integrity & experience for the following offices of EXECUTIVE Committee (EC) and Nomination Committee (NC).

For the Term 2019-2020 (upto March 31, 2020)

1. Vice-President (President Elect)
2. Nominations Committee (3 members)

For the Term 2019-2021 (upto March 31, 2021)

1. Hony. Treasurer
2. Regional Vice-President (Region 1): Delhi, Punjab, Haryana, Himachal Pradesh, Jammu & Kashmir, Uttar Pradesh, Uttarakhand & other areas of Northern India.
3. Regional Vice-President (Region 3): Gujarat, Madhya Pradesh & Rajasthan
4. Regional Vice-President (Region 5): Karnataka & Andhra Pradesh
5. Regional Vice President (Region 7) Tamil Nadu, Pondicherry, Andaman & Nicobar, Kerala And Lakshadweep
6. Divisional Chairperson (Division 1) – Systems
7. Divisional Chairperson (Division 3) – Applications
8. Divisional Chairperson (Division 5) –Education and Research

Note: All nominees and their proposers must refer to the eligibility conditions for the respective posts as outlined in the CSI Constitution and Byelaws.

The proposal for Nomination should be accompanied by (in prescribed format): which are also given in website www.csi-elections.org / www.csi-india.org..

1. Signed letter/ E-mail from at least two valid voting members proposing the Nominee.
2. A signed letter/ E-mail from the Nominee confirming:
   - 2.1 Acceptance to stand for election to the nominated office.
   - 2.2 Willingness to devote adequate time for the Society's work.
   - 2.3 Commitment to attend at least three ExecCom Meetings in a year (Not for Nominees to NC).
3. Two Passport size Photographs (printed) (3.5 cm X 4.5 cm) or softcopy of 413 x 531 pixels (300 dpi).
4. Statement of Intent on how the nominee intends to serve the Computer Society of India.
5. Bio-data in the suggested format (Please refer to formats for election)

Note-1: If the name of any Nominee appears for more than one Office, the Nominations Committee will be empowered to decide on the office for which he/she should contest. The NC will take into consideration any e-mail or signed written preferences submitted by the nominee received prior to the last date of nominations.

Note-2: Nominees will NOT be considered in the following cases:

(a) Nominees with pending dues to CSI or

(b) Nominees against whom Disciplinary action has been taken or

(c) Nominees with pending issues with the Disciplinary Committee.

Note-3: The nominee should be a valid member as on 31st March 2020. The nominee and the proposers must have paid their membership dues for the period for which he is contesting the election. e.g. nominee for Hony. Treasurer & other candidates whose term shall remain for two years must have paid membership dues for two years for which he is contesting i.e. till 2021.

Note-4: All election related notices will be published on the Official website for CSI Website Homepage i.e www.csi-elections.org/www.csi-india.org. The date of publishing election related notices on the CSI Website Homepage www.csi-elections.org/www.csi-india.org will be considered as the date of publication. As per Section 4.6.4, "The word mail includes e-mail and the word publication includes web publication".

The last date for receipt of nominations is November, 17, 2018 (5.00 pm)

The proposals must be sent to:

The Chairman, Nominations Committee (2018-19)
Computer Society of India
Samruddhi Venture Park, Unit No.3
4th. Floor, MIDC, Andheri (Eest)
Mumbai-400093
Ph.29261700
Email: nc2018-19@csi-india.org

With a copy to:
Mr. Anand Rao - Chairman, Nominations Committee
(E-mail: nanandrao@yahoo.com), Mob:9845400998.

S. No.	ACTIVITY	DATE/s
1	Start of Call for Nominations	2nd November, 2018
2.	Last Date for Receipt of Nominations	17th November, 2018 (5.00PM)
3.	Last date for Withdrawal of nominations	19th November, 2018 (5.00pm)
4.	Scrutiny and Finalization of Election Slate and its communication to ExecCom.	21st November, 2018
5.	Publication on CSI Elections Website with link of bio data and statement of Intent of candidates	10th December, 2018
6.	Email Posting of login and passwords to Members through Email.	31st December, 2018
7.	E-Voting starts (E-ballots)	1st January, 2019
8.	E-Voting Ends	31st January, 2019
9.	Declaration of Election Results on CSI Elections Website & communication to ExeCom	4th February, 2019
10.	Sending of Elections Results to the candidates & Registrar of Societies	5th February, 2019

Note: The dates may be changed by the Nominations Committee, if required – by suitable announcements on the CSI Website Homepage www.csi-elections.org/www.csi-india.org

All members are requested to register/update their latest e-mail Ids and mobile numbers well before the beginning of voting process by sending an email at nc2018-19@csi-india.org. **No change in email id or mobile would be permitted during the period of election**.

Election Code of Conduct

Canvassing (Both individually and in Group) during the Election Period directly or indirectly through Post/Email/Social Media/ SMS is not allowed / permitted. Action would be initiated against the Members/Non-members found involved in such an act of canvasing or otherwise, in accordance with latest version of Information Technology Act of the Govt. of India.

For more details on Elections, please visit www.csi-elections.org /www.csi-india.org

Nomination Committee (2018-19), CSI

N. Anand Rao
Chairman

Md. Shams Raza
Member

Sanjay Kumar Mohanty
Member

Call For Nominations

http://www.csi-elections.org

Format for Accepting Nomination

Date:__ / __/ ____

From:
Name :
CSI Membership No. :
Address :
City :
Mobile/ Phone No. :
Email :

To
The Chairman,
Adhoc Nominations Committee
C/o Computer Society of India-Delhi Chapter
No 32, DDA Market
QD Block, Muni Maya Ram Marg, Pitampura
Delhi-110034.

Sub: CSI Elections 2019-2020/2021 Nomination for the Post: _____

Dear Sir,

This is with reference to CSI elections.

1. I wish to inform you that my name has been proposed by some responsible eligible voting member of CSI having Membership No. ..... for the post of _____ for the term 2019-2020/2021.
2. My CSI membership no. is _____
3. My membership is valid upto (_____ date _____) OR I am a Life Member of CSI.
4. I accept my nomination for above post and stand for election to the nominated office.
5. I declare that I satisfy the eligibility requirements for the nominated office/post as specified in the CSI Constitution and byelaws.
6. I assure you that I will devote adequate time for the Society's work and attend at least 3 ExecCom meetings in a year. (Not for nominees of Nomination's Committee)
7. I declare that the information given above is true to the best of my knowledge and belief.
8. My photograph, bio-data and statement of intent are attached.

Thanking you,

Yours sincerely,

Signature of Nominee

Encl. : 1. Bio data with photograph and statement of intent in prescribed format.
2. Signed letter/E-mail from at least 2 valid voting members in good standing proposing the Nominee.

Call For Nominations

http://www.csi-elections.org

BIO DATA (format)

Date:__ / __/ ____

POST FOR WHICH NOMINATED: _____

1. Name :
2. CSI Membership No. :
3. Membership Type: Life/Annual/Institutional
4. CSI Membership since :
5. E-mail address :
6. Date of Birth (Age) :
7. Postal Address
8. Phone/ Mobile/ Fax Nos. :
9. Educational Qualifications :
10. Publications – relevant to the office being nominated for :
11. Contribution to the IT profession :
12. Contribution made to CSI :
13. Experience - relevant to the position nominated for :
14. Honours/ Professional Recognition :
15. Other Relevant Information :
16. In case of Nominees who are holding or have held an Elected post in CSI ExecCom or NC in the last 3 years :
    a. Positions held :
    b. Statements of Intent submitted for the above positions :
    c. Results achieved/ action taken :

Attach Photograph, Letter/Email of Nominee, Letter/Email of Proposer 1, and Letter/Email of Proposer 2

Signature of Nominee

Date :

**Give information in brief.(last 3 year

Computer Society of India

Format for Proposer

(Proposal from eligible voting member proposing the nominee)

(Two proposers are required for a valid Nomination)

Date:__ / __/ ____

From:
Name :
CSI Membership No. :
Address :
City :
Mobile/ Phone No. :
Email :

To
The Chairman,
Adhoc Nominations Committee
C/o Computer Society of India-Delhi Chapter
No 32, DDA Market
QD Block, Muni Maya Ram Marg, Pitampura
Delhi-110034.

Sub: Proposal for proposing the nominee : (name of the nominee) for CSI Elections 2019-2020/2021

Dear Sir,

This is with reference to CSI elections 2019-2020/2021.

1. I propose Mr./Prof./Dr./Mrs./Ms. _____ having CSI Membership No.:_____ as a candidate for the post of _____

2. I also certify that I have paid all my membership dues for the term of post for which I am proposing the candidate I have not nominated anyone else for the same post.

Thanking you,

Yours sincerely,

Signature
Name & CSI Membership Number of Proposer.

EXECUTIVE COMMITTEE

President (2018-19)

POST VACANT

(As the incumbent has been suspended)

Vice President (President Elect) (2018-19)

Prof. A. K. Nayak
Indian Institute of Business Management,
Budh Marg, Patna - 800 001
(T) 0612-2538809,
(M) 09431018581, 09386598581
(E) aknayak@iibm.in

Hony. Secretary (2018-2020)

Dr. Santosh Kumar Yadav
A-314, D.S.I.D.C. Flats
Paschim Puri
New Delhi-110063
(M) 098108 88851
(E) drskyadav@hotmail.com

Hon. Treasurer (2017-19)

Mr. Manas Ranjan Pattnaik
Plot No. N-24,
25 Chandaka Industrial Estate,
Patia, KIIT, Bhubaneswar
(M) 07873099999
(E) manas@anthemgt.com

President (2017-18)

Mr. Sanjay Mohapatra
Plot No. 5, CM 839/11, Sector 9
CDA, Market Nagar,
Cuttack - 753 014, Odisha.
(M) 91-9861010656
(E) smohapatra70@yahoo.co.in

REGIONAL VICE PRESIDENT

Region-I (2017-19)

Mr. Arvind Sharma
3/294, Vishwas Khand,
Gomati Nagar, Lucknow-226010. UP
(T) 522-4075496
(M) 9918653442 / 9415063442
(E) arvindsha@hotmail.com
a.arvind.sharma@gmail.com

Region-II (2018-20)

Dr. Jyotsna Kumar Mandal
University of Kalyani, Kalyani,
Nadia 741235, West Bengal
(T) 033-2580 9617
(M) 09434352214
(E) arvindsha@hotmail.com

Region-III (2017-19)

Dr. Vipin Tyagi
Dept of CSE
Jaypee University of Engg. and Tech.
Raghogarh, Guna - MP 473226
(T) 07544 - 267310-14 ext.134
(M) 09826268087
(E) dr.vipin.tyagi@gmail.com

Region-IV (2018-20)

Er. Nachindra K. Behera
Plot No : 223, Prachi Enclave-II
Chandrasekharpur
Bhubaneswar-751016, Odisha
(M) 9438838527
(E) nachindrabehera@gmail.com

REGIONAL VICE PRESIDENT

Region-V (2017-19)

Mr. Vishwas Bondade
No. 774, 2nd Stage, Indiranagar,
Bangalore 560038
(M) 09844058799
(E) vishwasbondade@gmail.com

Region-VI (2018-20)

Mr. Pradeep Rathi
E-401 Sky Anchorage
Panch Marg, Versova
Andheri (West),
Mumbai – 400061.
(M) 98202 93998 / 98200 19503
(E) pjrathi61@gmail.com

Region-VII (2017-19)

Dr. M. Sundaresan
Professor and Head,
Department of Information Technology,
Bharathiar University, Coimbatore - 641046, Tamil Nadu.
(M) 09443042340
(E) bu.sundaresan@gmail.com

DIVISION CHAIRPERSONS

Division-I (2017-19)

Mr. Apoorva Agha
8, Katra Road, Allahabad, UP - 211002
(M) 09415316183 / 08004905012
(E) apoorvaagha@hotmail.com
apoorvaagha@gmail.com

Division-II (2018-20)

Col. Balraj Anand
2/334, Guru Appartment
Sector 6, Dwarka, New Delhi
(M) 9811648050
(E) b_anand6@rediffmail.com

Division-III (2017-19)

Post Vacant
(As the incumbent has been suspended)

Division-IV (2018-20)

Prof. Vibhakar Mansotra
Department of Computer Science & IT
University of Jammu, Jammu (J&K)
(M) 9419103488
(E) vibhakar20@yahoo.co.in

Division-V (2017-19)

Dr. P. Kumar
Professor and Head
Department of Computer Science and Engineering,
Rajalakshmi Engineering College, Chennai – 602 105.
(M) 098405 73702
(E) pkumar_5@yahoo.com

NOMINATIONS COMMITTEE (2018-2019)

Mr. N. Anand Rao
2235,3rd Cross
Ragini Nilay, Banashankari
2nd Stage, Bangalore-560070
(M) 98454 00998
(E) nanandrao@yahoo.com

Md. Shams Raza
Ignou Programme Centre
St. Xavier School Campus
Gandhi Maidan, Patna – 800 001
(M) 94308 28918
(E) s_raza2000yahoo.com

Mr. Sanjay Kumar Mohanty
PACE, Padhuan Pada,
Proof Road, Balasore, Odisha
(M) 9437267606
(E) lmohantys@rediffmail.com

IPS COLLEGE OF TECH & MGMT., GWALIOR

In order to make the students to aware about the Cryptography, a workshop on the said subject was conducted by the CSI Gwalior Chapter in association and CSI Student chapter IPS College, on 24 Sept. 2018 at the Dept. of Computer & Engineering of IPS College. Mr. V. K. Arya (one of the subject expert and) life member of CSI, Delhi were gracious to deliver the speech. Shri Jayant Bhide, Immediate past Chairman CSI Gwalior (and Mr. Kailash Agrawal were present at this occasion. Er. P.S. Chauhan Principal IPS College along with their faculties of the department were taken all the pains to make this event successful. Students had taken very keen interest in this curious subject and it was enjoyed by all. This program has given a wide idea of Cryptography. Er. Arya praised all the students for the keen interest and the curiosity shown by them. Principle Er. Chauhan praised everyone and express his happiness to see the success of the event. Prof. Karunendra Verma, has given the vote of thanks. Overall it was a very good program.

Call for Paper for CSI Journal of Computing

(e-ISSN: 2277-7091)

Original Research Papers are invited for the CSI Journal of Computing, published on line quarterly (e-ISSN: 2277- 7091) by the Computer Society of India (CSI). The Journal of Computing, offers good visibility of online research content on computer science theory, Languages & Systems, Databases, Internet Computing, Software Engineering and Applications. The journal also covers all aspects of Computational intelligence, Communications and Analytics in computer science and engineering. Journal of Computing intended for publication of truly original papers of interest to a wide audience in Computer Science, Information Technology and boundary areas between these and other fields.

The articles must be written using APA style in two columns format. The article should be typed, double-spaced on standard-sized (8.5" x 11") with 1" margins on all sides using 12 pt. Times New Roman font and 8-12 pages in length. The standard international policy regarding similarity with existing articles will be followed prior to publication of articles. The paper is to be sent to **Prof. (Dr.) J. K. Mandal**, Editor-in-Chief, CSI Journal of Computing (**csi.journal@csi-india.org**) with a copy to the Publisher Prof. A K Nayak in the email id : aknayak@iibm.in.

Prof. A K Nayak
Publisher

REGION-I

Amity University, Noida

30-8-2018 – Competition (Decrypton 4.0)

REGION-II

Techno International, Batanagar

10-10-2018 – Student Branch Inauguration

REGION-II

Budge Budge Institute of Technology, Kolkata

11-9-2018 – Student Branch Inauguration

REGION-III

Devang Patel Inst. of Advance Tech. and Research, Anand

1-9-2018 - Workshop on Introduction to Deep Learning and its applications on NVIDIA Graphics Card

REGION-III

Devang Patel Inst. of Advance Tech. and Research, Anand

29-9-2018 – Hands-on workshop on MATLAB Programming

REGION-IV

Shri Shankaracharya Technical Campus, Bhilai

8-10-2018 to 13-10-2018 - Workshop on Data analysis & Machine Learning Using Python

REGION-IV

Shri Shankaracharya Technical Campus, Bhilai

8-10-2018 – Student Branch Inauguration

Region-V

Lakireddy Bali Reddy College of Engineering, Mylavaram

7-9-2018 - One day workshop on 5G Technologies (A New Era in Communication) by Dr Koteswararao Kondepu

REGION-V

Sasi Institute of Technology & Engineering, Tadepalligudem

19-9-2018 - Guest Lecture on Recent trends in Software Industry and how Industry works

26-9-2018 - CSI Swearing Ceremony of New Office Bearers & Seminar on Block Chain Architecture and Design

G Pullaiah College of Engineering & Technology, Kurnool

4-10-2018 & 5-10-2018 - Workshop on **Machine Learning**

PVP Siddhartha Institute of Technology, Vijayawada

6-10-2018 – FDP on Self-Motivation Towards R & D Activities by Prof P Thrimurthy, Fellow and Past President, CSI

NMAM Institute of Technology, NITTE

12-9-2018 - INQUIZITIVE

19-9-2018 - Motivational talk Learnings from the Optipace Journey

Vidya Jyothi Institute of Technology, Hyderabad

28-9-2018 & 29-9-2018 - Hands on Workshop on Augmented Reality and Virtual Reality

6-10-2018 & 7-10-2018 - Hands-on Workshop on Machine Learning

REGION-V

New Horizon College of Engineering, Bangalore

31-8-2018 – Dr Prakash, Past Chairman CSI Bangalore Chapter inaugurated Student Branch

6-10-2018 - Workshop on Python for Data Science and Machine Learning

REGION-VII

Velalar College of Engineering & Technology, Erode

26-9-2018 – Dr. M. Sundaresan, RVP-VII inaugurated the Student Branch

S.A. Engineering College, Chennai

7-9-2018 - SA(HA) Ckathon Coding Contest

Sethu Institute of Technology, Kariapatti

11-7-2018 - Guest Lecture on Latest Technology in Cloud Computing and Big Data with Intelligence

28-9-2018 & 29-9-2018 - Technical seminars on Cloud Computing

Second International Conference

on

Advanced Computational and Communication Paradigms (ICACCP-2019)

ICACCP'19 PROJECT SHOWCASE

Date: February 25-28, 2019

Organized by:

Department of Computer Science and Engineering

Sikkim Manipal Institute of Technology

OBJECTIVES OF THE SHOWCASE

- To create a culture of innovation in an educational institute
- To develop entrepreneurial mindset among students
- To make students aware of latest industry trends
- To make students aware of government support available to them to carry forward their projects as entrepreneurial ventures
- To connect with likeminded innovators for further enhancement of the project

THEME OF SHOWCASE

- Computational and Communication Paradigms
- State- of-the Art Design
- Technology and Application
- Science projects but not limited to ICACCP-2019 theme

ATTRACTIVE PRIZE FOR BEST THREE PROJECTS

CONTACTS:

Prof. Sourav Paul (9748445834)
Prof. Chinmoy Kar (9874365360)
Prof. S. K. Mishra (8927257899)
Dr. Aniruddha Dey (8240761221)
Email: icaccp2019@smit.smu.edu.in

Participant	Registration Fee
Indian	₹ 2,000
Foreign	USD 100

Submission through Easy Chair:
https://easychair.org/conferences/?conf=icaccp2019

IMPORTANT DATES:

Submission Deadline: 31st Dec 2018
Acceptance Notification: 15th Jan 2019
Final Submission Due: 31st Jan 2019
For more details visit:
http://www.icaccpa.in/

Registered with Registrar of News Papers for India - RNI 31668/1978
Regd. No. MCN/222/2018-2020
Posting Date: 10 & 11 every month. Posted at Patrika Channel Mumbai-I
Date of Publication: 10th of every month

If undelivered return to :
Samruddhi Venture Park, Unit No.3,
4th floor, MIDC, Andheri (E). Mumbai-400 093

INDIACOM-2019

(IEEE Conference ID: 46181 | SCOPUS Indexed)

13th INDIACom; 2019 6th IEEE International Conference on

"Computing for Sustainable Global Development"

(13th – 15th March, 2019)

Organized by

Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM), New Delhi

Technically Sponsored by

IEEE Delhi Section

Supported by

Computer Society of India (CSI), Divisions – II, IV and Delhi Chapter, Institutions of Electronics and Telecommunications Engineers (IETE), Delhi Centre, Indian Society for Technical Education (ISTE), Delhi Section and Guru Gobind Singh Indraprastha University (GGSIPU), New Delhi

Paper Submission Deadline: 31st October, 2018 [No Further Extension]

Paper submission Link: http://bvicam.ac.in/indiacom/submitPaper.asp

Conference Website: http://bvicam.ac.in/indiacom/

Announcement and Call for Papers

INDIACom-2019 is aimed to invite original research papers in the field of, primarily, Computer Science and Information Technology and, generally, all interdisciplinary streams of Engineering Sciences, having central focus on sustainable computing applications, which may be of use in enhancing the quality of human life and contribute effectively to realize the nations' vision of sustainable inclusive development using Computing. INDIACom-2019 will be an amalgamation of four different Tracks organized parallel to each other, in addition to the 05th International Workshop on Information Engineering and Management (IWIEM-20197) and few theme based Special Sessions, as listed below:-

- **Track #1: Sustainable Computing**
- **Track #2: High Performance Computing**
- **Track #3: High Speed Networking & Information Security**
- **Track #4: Software Engineering & Emerging Technologies**
- **Track #5: Theme Based Special Sessions**

Instruction for Authors

Authors from across different parts of the world are invited to submit their papers. Authors should submit their papers online at http://www.bvicam.ac.in/indiacom/loginReqSubmitPaper.asp. New authors should first sign up and create an account on http://www.bvicam.ac.in/indiacom/addMember.asp to log in and submit paper. Only electronic submissions will be considered. Paper submission, as E-Mail attachment, will not be considered.

Important Dates

Submission of Full Length Paper	31st October, 2018	Paper Acceptance Notification	15th January, 2019
Submission of Camera Ready Copy (CRC) of the Paper	29th January, 2019	**Registration Deadline** (for inclusion of Paper in the Proceedings)	29th January, 2019

Accepted Papers will be published in IEEE Xplore, which is indexed with world's leading **Abstracting & Indexing (A&I)** databases, including ISI, SCOPUS, DBLP, EI-Compendex, INSPEC, Google Scholar, etc. Further details are available at www.bvicam.ac.in/indiacom. All correspondences, related to INDIACom-2019, must be addressed to:

Prof. M N Hoda
General Chair, INDIACom-2019
Director, Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM)
A-4, Paschim Vihar, Rohtak Road, New Delhi-110063 (INDIA)
E-mails: conference@bvicam.ac.in, indiacom2019@gmail.com
Tel.: 011-25275055 TeleFax: 011-25255056, 09212022066 (Mobile)