



CSI Communications

Knowledge Digest for IT Community

VOLUME NO. 43 | ISSUE NO. 2 | MAY 2019

₹ 50/-

DATA SCIENCE AND ANALYTICS

COVER STORY

Machine Learning and Data Analytics: Need of the Hour **06**

TECHNICAL TRENDS

Evolution of Data Analytics in the IOT Era... **16**

RESEARCH FRONT

Monetize Dormant Data: 3 Steps for C-Suite Executives to Integrate Analytics into Decision Making **23**

ARTICLE

Managing Credit Risk in Banking Application using Data Analytics – A Bird View **29**

SECURITY CORNER

Security Aspects in E-Commerce **37**

Know Your CSI

Executive Committee (2018-19/20)



President
Prof. A K Nayak
(E) aknayak@iibm.in



Vice President cum President Elect
Mr. R K Vyasa
(E) rkvyasa05@gmail.com



Hon. Secretary
Dr. Santosh Kumar Yadav
(E) secretary@csi-india.org



Hon. Treasurer
Dr. Durgesh Mishra
(E) drdurgeshmishra@gmail.com



Immd. Past President (2017-18)
Mr. Sanjay Mohapatra
(E) smohapatra70@yahoo.co.in

Regional Vice-Presidents



Region-I (2019-21)
Mr. Arvind Sharma
(E) rvp1@csi-india.org



Region-II (2019-21)
Md. Shams Raza
(E) s_raza2000@yahoo.com



Region-III (2019-21)
Mr. Jayant Binde
(E) jayantbhide2000@gmail.com



Region-IV (2018-20)
Er. Nachindra K Behera
(E) nachindrabehera@gmail.com



Region-V (2019-21)
Mr. M. S. Prasadbabu
(E) profmspbabu@gmail.com



Region-VI (2018-20)
Mr. Pradeep Rath
(E) prathid61@gmail.com



Region-VII (2019-21)
Prof. Dr. M Sundaresan
(E) bu.sundaresan@gmail.com



Region-VIII (2019-21)
Mr. Konatham Mohan Raidu
(E) kmraidu@informaticsindia.com

Division Chairpersons



Division-I (2019-21)
Mr. Manas Ranjan Pattnaik
(E) manas@anthemgmt.com



Division-II (2018-20)
Col. Balraj Anand
(E) b_anand6@rediffmail.com



Division-III (2019-21)
Prof. Suresh Chand Tyagi
(E) sct_35_2000@yahoo.com
sctyagi1963@gmail.com



Division-IV (2018-20)
Prof. Vibhakar Mansotra
(E) vibhakar20@yahoo.co.in



Division-V (2019-21)
Dr. Subhash Chandra Yadav
(E) scy123@gmail.com

Nomination Committee (2019-2021)



Arvind M Nayak
(E) amnayak@hotmail.com



Mr. Deepak Sharma
(E) dsharmaji@gmail.com



Dr. R R Deshmukh
(E) rrdeshmukh.csit@bamu.ac.in

Dear All,

Thank you for your affiliation with us.

To show our appreciation we have come up with
Exclusive Summer Discount for life membership.

Life membership Fees (including GST)- ₹ 11,800/-

After availing 15% discount- ₹ 10,030/-

This discount will give you a lot of savings. Members are and will always remain our assets. We request you all to spread the message of this exclusive offer among your circle to avail the discount before the expiry date.

It is an honour for us to count your great contribution for membership development of the society and we hope to get the same from you in future also to get new enrolments from your known circle.

In case of any queries please feel free to write to us at sonali@csi-india.org / swapnil@csi-india.org. So Hurry up and grab this opportunity. Thanking You in appreciation & looking forward to your kind cooperation.

With Best Regards

Dr. S K Yadav
Hony. Secretary (CSI)
(2018-2020)

**CSI SUMMER
OFFER**

15th May - 15th July 2019





Chief Editor
S S AGRAWAL
KIIT Group, Gurgaon

Guest Editor
PROF. NASIB SINGH GILL
Sr. Professor of Department of
Computer Science & Applications,
M. D. University, Rohtak
Email: nasibsgill@gmail.com

Published by
AKSHAYA KUMAR NAYAK
For Computer Society of India

Editorial Board:
Arun B Samaddar, NIT, Sikkim
Bhabani Shankar Prasad Mishra,
KIIT University, Bhubaneswar
Debajyoti Mukhopadhyay, MIT, Pune
J Yogapriya, Kongunadu Engg. College, Trichy
M Sasikumar, CDAC, Mumbai,
R Subburaj, SRM University, Chennai
R K Samanta, Siliguri Inst. of Tech., West Bengal
R N Behera, NIC, Bhubaneswar
Sudhakar A M, University of Mysore
Sunil Pandey, ITS, Ghaziabad
Shailesh K Srivastava, NIC, Patna
Vishal Mehrotra, TCS

Design, Print and Dispatch by
GP OFFSET PVT. LTD.

Please note:
CSI Communications is published by Computer Society of India, a non-profit organization. Views and opinions expressed in the CSI Communications are those of individual authors, contributors and advertisers and they may differ from policies and official statements of CSI. These should not be construed as legal or professional advice. The CSI, the publisher, the editors and the contributors are not responsible for any decisions taken by readers on the basis of these views and opinions.
Although every care is being taken to ensure genuineness of the writings in this publication, CSI Communications does not attest to the originality of the respective authors' content.
© 2012 CSI. All rights reserved.
Instructors are permitted to photocopy isolated articles for non-commercial classroom use without fee. For any other copying, reprint or republication, permission must be obtained in writing from the Society. Copying for other than personal use or internal reference, or of articles or columns not owned by the Society without explicit permission of the Society or the copyright owner is strictly prohibited.

CONTENTS

Cover Story

Machine Learning and Data Analytics: Need of the Hour 6
Amit Sagu, Tarika Verma and Nasib S. Gill

An impact of Big Data and Crowdsourcing Analytics in business process 9
S. Balakrishnan, Jayasudha Subburaj and J. P. Ananth

Machine Learning and Analytics 12
Nasib S. Gill and D. Seehrawat

Role of Data Analytics in Health Care Sector – A Research View 14
T. Ramesh and V. Santhi

Technical Trends

Evolution of Data Analytics in the IOT Era... 16
S. Shenbagavadivu, M. Senthil Kumar and B. Chidambara Rajan

Research front

Monetize Dormant Data: 3 Steps for C-Suite Executives to Integrate Analytics into Decision Making 23
Binesh Nair

A Framework for Software Fault Prediction using Decision Tree Based Machine Learning 26
Preeti Gulia and Palak

Articles

Managing Credit Risk in Banking Application using Data Analytics – A Bird View 29
Archana Sasi

Handwritten Digits Recognition using Convolutional Neural Network 31
Amit Choudhary and Savita Ahlawat

Role and Impact of Big Data Analytics in Health Care Systems 34
Mamoon Rashid, Harjeet Singh, Vishal Goyal, Kulvinder Singh and Neeraj Mogla

Security Corner

Security Aspects in E-Commerce 37
Gopal Krishna, Aayush Aman and Tushar Anand

PLUS

CSI Summer Offer	2nd Cover
CSI Publications Team	15
CSI Executive Committee	18
Summer Offer for Life Membership	20
Congratulations: Prof. Sinha appointed as International Development Consultant	30
About the Guest Editor: Prof. Nasib Singh Gill	33
CSI Region – IV Regional Student Convention - A Report	40
Seminar conducted on 'D-Link Academy Campus Connect Program'	41
Mahatma Gandhi Mission's College of Engg. and Tech., Kamothe, Navi Mumbai	41
CSI Regional Student Convention- Region-III "KAUSHAL" -2019	42
Call for Regional & State CSI Student Coordinator/s - 2019-2020	44
Student Branches Inauguration Report	45
From CSI Chapters & Divisions	46
Call for Paper for CSI Journal of Computing	47
From CSI Student Branches	48
CSI 53rd Annual Convention	Back Cover

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: S. S. Agrawal

Editorial



Dear Fellow CSI Members,

"Things get done only if the data we gather can inform and inspire those in a position to make [a] difference."

- *Mike Schmoker*

"Without big data analytics, companies are blind and deaf, wandering out onto the web like deer on a freeway." - *Geoffrey Moore*

After getting familiar with era of E-commerce in April – 2019 issue, we are now focusing on "Data science and Analytics" - an important, interdisciplinary area of research and application. The current issue flags the information of articles related with Data science and Analytics.

The need of data analytics has been focused by Amit Sagu and Nasib S. Gill in cover story entitled "Machine Learning and Data Analytics: Need of the Hour". An impact of Big Data and Crowdsourcing Analytics in Business Process has also been explained by S. Balakrishnan, Jayasudha Subburaj and J. P. Ananth in another cover story, relationship between machine learning and analytics given in cover story by Nasib S. Gill and D. Sehwat. A research view on role of data analytics in health care sector has been explained by T. Ramesh and V. Santhi in last cover story.

The Technical trends is showcased with article "Evolution of Data Analytics in the IOT Era..." by S. Shenbagavadivu, M. Senthil Kumar and B. Chidambaram Rajan.

In Research Front we have contributions by Binesh Nairas, "Monetize Dormant Data: 3 Steps for C-Suite Executives to Integrate Analytics into Decision Making" and "A Framework for Software Fault Prediction using Decision Tree Based Machine Learning" by Preeti Gulia and Palak.

Other articles in this issue are,

- "Managing Credit Risk in Banking Application using Data Analytics – A Bird View" by Archana Sasi
- "Handwritten Digits Recognition using Convolutional Neural Network" by Amit Choudhary and Savita Ahlawat
- "Role and Impact of Big Data Analytics in Health Care Systems" by Mamoon Rashid, Harjeet Singh, Vishal Goyal, Kulvinder Singh and Neeraj Mogla

In Security Corner, this issue have article on "Security Aspects in E-Commerce" written by authors Gopal Krishna, Aayush Aman and Tushar Anand.

The details about new members of Executive Committee for the term (2019-20/21) and publication team has been given.

This issue also gives detailed information about reports on the CSI Region-IV Regional Student Convention that was organized by CSI Student Branch – GIET University, Gunupur in collaboration with CSI - Gunupur Chapter on 2019 on "Artificial Intelligence and Machine Learning" during 24-25 February 2019. This issue is also includes a report on seminar conducted on 'D-Link Academy Campus Connect Program' organized by Universal College of Engineering in association with CSI, two days' workshop on "Android Application Development" conducted by the IT Department of Mahatma Gandhi Mission's College of Engg. and Tech., Kamothe, Navi Mumbai on 5th & 6th April 2019, a Regional Student Convention-Region III "KAUSHAL-19" organized on 15th - 16th March 2019 at Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore hosted by CSI, Indore Chapter and CSI-Student Branch, SVV Indore. This issue also contains two student branches inauguration reports opened in last months, first at University of Engg. & Mgmt. (UEM), Kolkata, West Bengal (Region-II) on 13th April 2019 and RWMCT's Dnyanshree Institute of Engineering & Technology, Satara, Maharashtra (Region-VI) on 19th March 2019. CSI activity reports from chapters, divisions and student branches are also included.

We are thankful to all contributors and look forward to receive your valuable articles in future also.

We express our gratitude to the CSI publication committee, editorial board, authors and reviewers for their great contribution and support to these issues.

Our special thanks to Prof. A. K. Nayak, President, CSI for his constant encouragement, support and enthusiasm in this publication of May 2019 issue.

With kind regards,

Prof. (Dr.) S. S. Agrawal

Director General KIIT & Emeritus Scientist (CSIR)



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



Dr. Nasib Singh Gill
Guest Editor

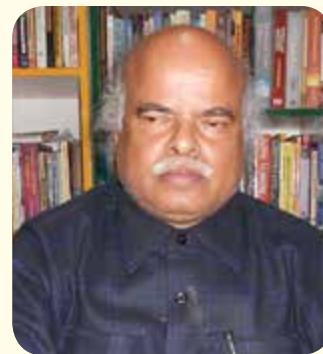


President's Desk

From : President, Computer Society of India

Date : 01 May, 2019

Email : president@csi-india.org / Cell : (91) 82106 93239



The theme of this issue of CSI Communication Data Science and 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 months of March & April have witnessed more than 100 events which were mentioned in last April issue & this issue of CSI Communication. In the coming months the quantum of such activities shall be in progress & such events will be regularly covered in CSI communication.

I request the Regional Vice Presidents, Divisional Chairmen, 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 its 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. I am happy to note that two Student Branches were inaugurated in last months, one is at University of Engineering and Management, Kolkata & another at Dayanshree institute of Engineering & Technology at Satara, Maharashtra. I congratulate the management of both the Institutes in particular & all the concerned students and faculty members in general for their pioneering efforts in establishment of these two new student branches. I also congratulate CSI Kolkata Chapter & Mr. Pradip Rathi, Regional Vice President, Region-VI for the same.

Let us join with me to welcome & congratulate our new publication team which shall be headed by the Eminent Scientist, Academic Administrator & great Visionary Dr. D. D. Sharma, Fellow & Life Time Achievement Awardee of CSI as the Chairman Publication Committee along with the Emeritus Scientist (CSIR) Dr. S. S. Agrawal, Fellow & Life Time Achievement Awardee of CSI as Chief Editor of CSI Communications, Ms. Ritika Wasson of Bharati Vidyapeeth Institute of Computer Application &

Management. New Delhi as Editor, DR. R. R. Deshmukh, Professor of Computer Science & Information Technology of Baba Saheb Ambedkar Marathawada University, Maharashtra as Chief Editor of CSI Journal of Computing, Dr. Brojo Kishore Mishra of GIET University, Gunupur, Odisha as Editor of SCI Journal of Computing and Dr. Vishal Jain of Bharati Vidyapeeth Institute of Computer Applications & Management, New Delhi as Editor of CSI Adhyayan.

CSI has invited applications for Regional Student Convenor & State Student Convenor in this issue of CSI Communication. I request the life members having significant contribution for the development of CSI student Community should come forward to apply for the same within the prescribed time.

CSI has also announced summer offer for membership discount of 15% from May 2019 to July 2019. I request all Hon'ble members to motivate to their known people to utilise this opportunity for more & more enrolment.

Though CSI is having good numbers of Chapters, Institutional Members, Students Branches still then its 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.

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.

With warm regards,

Prof. Akshaya Nayak
President, CSI

Machine Learning and Data Analytics: Need of the Hour

► Amit Sagu, Tarika Verma and Nasib S. Gill

Department of Computer Science & Applications, Maharshi Dayanand University, Rohtak, Haryana (India). Email: nasibsgill@gmail.com

1. Overview

Machine learning (ML) and data analytics are subset of Artificial intelligence (AI). AI or sometime referred as Machine Intelligence which mimic human intelligence. The ambition of AI is to make intelligent machine which can take decision depend on given state. The concept of AI is not new it has been introduced a couple of years ago but due to insufficient device which could not have able to fabricate or process data. Today we have Integrated Circuit (IC) which can store data exponentially and we are capable to make machines learn.

2. Machine Learning

Machine Learning is an application of AI which give system ability to learn and improve from learning experience without giving explicitly instruction [1].

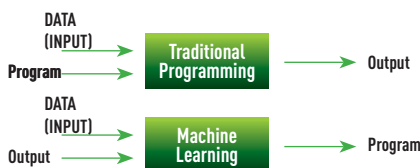


Fig. 1: Machine learning vs traditional programming

ML is similar to how we train any child. It is all about trained the machine first and then machine predict the outcome from previous learning experience. ML is different from traditional programming, in traditional Programming we feed input (data) and program (logic) to machine to get the output. But in ML we feed input data and output data to make its own program it is known as training [2]. Machine learning mainly classifies into **supervised learning** and **unsupervised learning**.

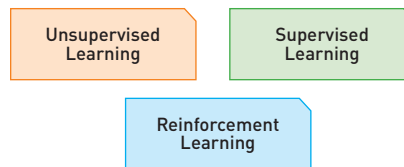


Fig. 2: classification of ML

1.1 Supervised Learning

As name evident supervisor or teacher to teach the machine first with known answer

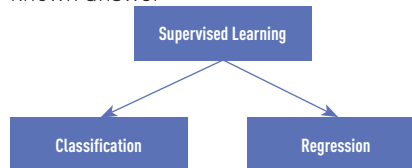


Fig. 3: types of supervised learning

Most of the machine learning algorithm are supervised. In it we train the machine with labeled data set.

We have input variable (x) and output variable (y), we use algorithm which learn mapping function from input data to output data [4].

$$Y=f(X)$$

In above equation 'f' is a function which take input variable (x) and mapping it to output (y). The goal is to approximate mapping so that when we have new input data (x) then we can predict the output variable (y).

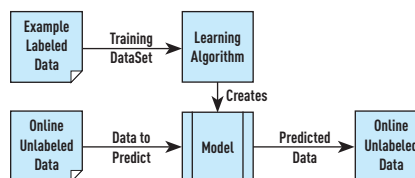


Fig. 4: supervised learning block diagram

Supervised machine learning can be classified into **Classification** and **Regression**.

2.1.1 Classification

It is supervised learning where output is a discrete value. It classify the output value in to homogeneous group. Its aim to predict discrete value belonging to specific class or group. It can be either multi class or binary class classification. For instance Gmail classify incoming email to spam, social, personal, organizational etc. this is multi class classification if we talk about binary class classification then Gmail classify incoming email to Spam or Non-spam. Fig (5) shows binary classification which have two different pattern.

Some other examples are:

- "Today will be rain or not",
- "is this picture flower or not".

Classification supervised learning is not suitable for data which have real values like weight or distance etc.

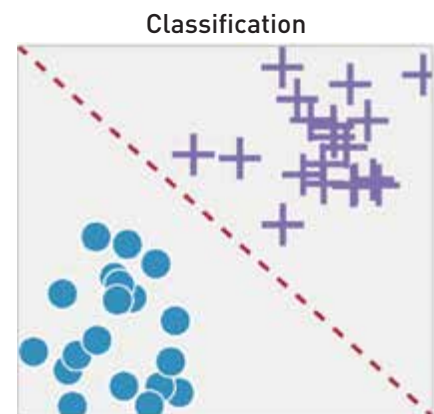


Fig. 5: Classification Supervised Learning

Example of classification task:

- Predict the picture of fruits.
- Predict the gender based on their appearance
- Predict the handwriting of gender.

2.1.2 Regression

Regression model is used to predict the continuous value like predicting the salary of employee based on the age.

For example which of the following a regression task?

- Predict the age of the person
- Predict the nationality of the person
- Predict whether it rain today
- What would be price of house after one year

Solution: predicting the age of person is regression because age is real of continuous value, "predict the nationality" is not regression because it is categorical, "predict whether it rain today" can we answered in yes or no, the price of house is also real value.

Regression

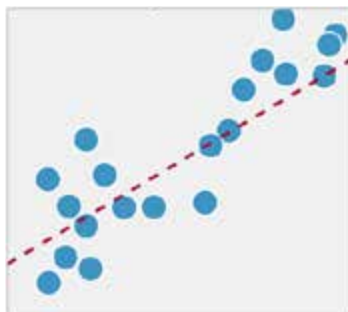


Fig. 6: regression supervised learning

2.2 Unsupervised Learning

Unsupervised learning is where we have only input data (X) and not having output variable (Y).

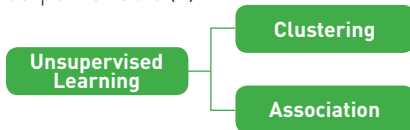


Fig. 7: unsupervised learning classification

Unlike supervised learning there is no teacher or supervisor to trained machine and learning algorithm act on information without any guidance. The main task of machine is group the unsorted data according to their pattern, similarities etc. [5]. For instance it is given a picture of fruits which have not seen ever, machine has no idea about

features of fruits. It cannot categorized into apple, orange or grapes etc., but it can categorized them according to their structure, with the fact that all apple have almost same shape. Machine can easily categorized them into different groups in spite of that machine didn't learn anything before.

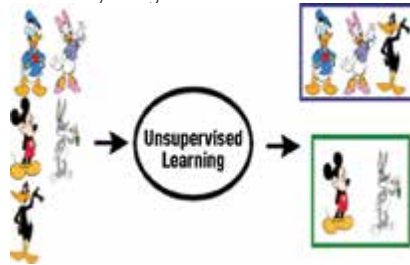


Fig. 8: unsupervised learning

2.2.1 Clustering

In it data is divided into several groups with similar patterns [6]. Clustering make homogenous data group and each group datasets are similar to another. Some of the most common clustering algorithms:

- K-mean
- Hierarchical Clustering
- Density based scan Clustering
- Gaussian Clustering Model

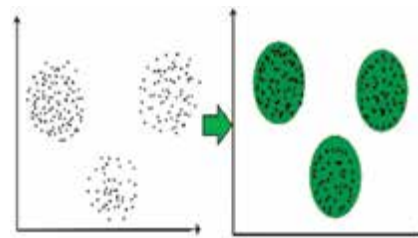


Fig. 9: Clustering unsupervised learning

2.2.2 Association

It is rule based machine learning technique which find out the relation between parameter of large datasets. Shopping store uses algorithm to find the relationship between products. For example {Egg + Bread} → Ketchup, means they put down ketchup near bread and eggs so that customer would purchase it even he was not intended. If model trained well then it can be used to increase the sale of products.

2.3 Reinforcement Learning

The main idea behind reinforcement learning is that a software agent interact

with environment and get reward for performing action.

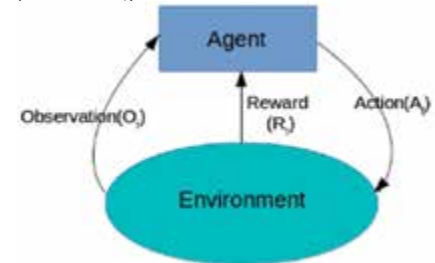


Fig. 10: reinforcement learning

Unlike supervised learning there is no training labeled datasets. Agent learn from its previous experiences. Reinforcement learning is work same as human learning characteristics. For example human baby see a fireplace he approach it and feel warmed (get reward +1), it get reward cause its feel positive and warm, but when it get too close it burns his hand (reward -1) . Reinforcement is a computational approach learning from action.

3. Data Analytics

Data Analytics is process of examining the data in order to draw inference through the information they contain [3].

Type of Data Analytics

- Predictive Analytics
- Descriptive Analytics
- Prescriptive Analytics

Predictive Analytics is a process which uses statistic model and machine learning to predict the outcome based on current and past events. Descriptive Analytics is simple form of Analytics which aggregate the whole data and provide insights. Prescriptive Analytics uses business rules, machine learning and statistics model in order to recommend the best course of action. Neural Network is a building block of Data Analytics. Neural Network is system of hardware and software that mimic the central nervous system of human.

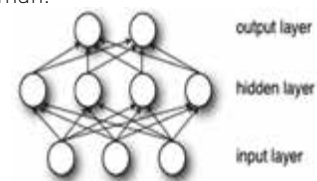


Fig. 11: Neural Network model

A neural network is a function which learn from expected output for a given input training datasets. As we can see in fig there are multiple hidden layer nodes which works as subsequent input. Each circle node have their own knowledge set

4. Conclusion

Machine learning is the hottest topic in IT. AI, Machine Learning, Automation are changing the shaping of future work. Semi-supervised learning is also introduced which combination of both supervised and unsupervised

is learning which could have labeled and unlabeled data. Data Analytics and its neural models are very popular in Machine Learning.

5. References

- [1] <https://www.coursera.org/learn/machine-learning>
- [2] <https://www.expertsystem.com/machine-learning-definition/>
- [3] Ivy Pro School, "6 Companies that use Machine Learning extensively", Available: <http://ivyproschool.com/blog/2017/08/24/6-companies-that-use-machine-learning-extensively/>
- [4] <https://towardsdatascience.com/>

supervised-machine-learning-classification

- [5] <https://www.mathworks.com/discovery/unsupervised-learning>
- [6] <https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/>
- [7] Vas3k blog, "Machine Learning for Everyone", Available: https://vas3k.com/blog/machine_learning/
- [8] Zhiqiang Ge, Zhihuan Song et al.; "Data Mining and Analytics in the Process Industry: The Role of Machine Learning", Access IEEE, vol. 5, pp. 20590 - 20616, 2017

About the Authors



Dr. Nasib Singh Gill is at present senior most Professor of Department of Computer Science & Applications, M. D. University, Rohtak, India and is working in the Department since 1990. He earned his Doctorate in Computer Science in the year 1996 and carried out his Post-Doctoral research at Brunel University, West London during 2001-2002. He is a recipient of Commonwealth Fellowship Award of British Government for the Year 2001. Besides, he also has earned his MBA degree. He has published more than 245 research papers in reputed National & International Journals, Conference Proceedings, Bulletins, Edited Books, and Newspapers. He has authored seven books. He is a Senior Member of IACSIT as well as a fellow of several professional bodies including IETE and CSI. He has been serving as Editorial Board Member, Guest Editor, Reviewer of International/National Journals and a Member of Technical Committee of several International/National Conferences. He has guided so far 9 Ph.D. scholars as well as guiding about 7 more scholars presently in the areas – IoT, Machine Learning, Information and Network Security, Computer Networks, Measurement of Component-based Systems, Complexity of Software Systems, Decision Trees, Component-based Testing, Data mining & Data warehousing, and NLP.



Mr. Amit Sagu has passed M.Sc. in 2016 in Computer Science and Applications from Department of Computer Science & Applications, Kurukshetra University Kurukshetra, India. He has also worked as Assistant Professor at DAV Centenary College, Faridabad. He is currently pursuing Ph.D. in Computer Science at M. D. University, Rohtak. His research interests include IoT, Machine Learning, Big Data Analytics and Data Mining.



Ms. Tarika Verma passed her B.Tech. and M.Tech. in Computer Science and Engineering from UIET, Maharshi Dayanand University, Rohtak, India. She has also worked as Assistant Professor at AIJHM College, Rohtak. She is currently pursuing Ph.D. in Computer Science at M. D. University, Rohtak. Her research interests include IoT, Machine Learning, Big Data Analytics and Data Mining.

KIND ATTENTION !

Prospective Contributors of CSI Communications

Fourth Coming Issues : June 2019 : Emotional Intelligence

Please note that Cover Theme for **June 2019 issue is Emotional Intelligence**. 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 May, 2019.

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 Editor **Ritika Wasan**, Associate Professor Bharati Vidyapeeth's Inst. of Computer Applications and Management (BVICAM) in the email id : **rit_2282@yahoo.co.in**

Issued on the behalf of Editorial Board, CSI Communications.

Prof. (Dr.) S S Agrawal
Chief Editor

An impact of Big Data and Crowdsourcing Analytics in business process

► S. Balakrishnan

Professor, Department of CS and Engg.,
Sri Krishna College of Engineering and
Technology, Coimbatore, Tamilnadu, India.

► Jayasudha Subburaj

Professor, Dept. of Master of Comp. Applications,
Sri Krishna College of Engineering and Technology,
Coimbatore – 641008, Tamil Nadu, India.

► J. P. Ananth

Professor, Dept. of CS and Engineering,
Sri Krishna College of Engineering and
Technology, Coimbatore, Tamilnadu, India.

Overview

As many big data organizations increase gigantic interests in big data to catch business bits of knowledge by scrambling to utilize information researchers, information designers and information investigators. Big data crowdsourcing can enhance associations speculation plans. For Silicon Valley new companies propelling a major information stage, the most ideal approach to diminish costs is to pay telecommuters with the goal that they can disseminate assignments to individuals who have web get to anyplace on the planet. Huge Data holds the guarantee of changing how organizations and individuals take care of true issues and Crowdsourcing assumes a crucial job in overseeing enormous information. In this article, we will talk about how publicly supporting huge information can reform business forms.

1. Introduction

Crowdsourcing, a combination of "crowd" and "outsourcing" first authored by Wired magazine in 2005 and energized by the Internet, is an amazing sourcing model that use the profundity of experience and thoughts of an open gathering instead of an

associations claim representatives. Matt H. Evans points out the importance of CrowdSourcing and he said that "Crowdsourcing taps into the global world of ideas, helping companies work through a rapid design process. You outsource to large crowds in an effort to make sure your products or services are right." The upsides of utilizing crowdsourcing are professed to incorporate improved costs, speed, quality, adaptability, versatility, or assorted variety. It has been utilized by new companies, expansive partnerships, non-benefit associations, and to make normal products. crowdsourcing is a case of ICT marvel based collaboration, collection, cooperation, agreement, and imagination. It is another method for doing work, where if the conditions are correct, the group can outflank singular specialists. Geologically scattered individuals associated by web can cooperate to deliver strategies and results that are worthy to most.

The key elements of crowdsourcing are as per the following:

1. An association that has an errand it needs performed,
2. A people group (crowd) that is happy to play out the errand willfully,

3. An ICT situation that enables the work to occur and the network to collaborate with the association,
4. Shared advantage for the association and the network.

1.1 Big Data Analytics

Big Data Analytics (BDA) is one of the most envisage fields in the present era after cloud computing. Big business houses and internet giants are busy to explore the benefits of BDA and have implemented the concept terrifically in the last decade to bring a great revolution in the field of data search, online retailing, digital marketing, web mining, social networking, community site growth and much more. Automatic data analysis techniques (for example AI) are frequently considered as principle segments of information investigation. Data analysis is intensely work concentrated. Manual handling rules a vast bit of information investigation process.

Crowdsourcing is from numerous points of view identified with huge information. Big data is an "expansive term for informational indexes so vast or complex that customary information handling applications are deficient". Big data was brought into the world with the

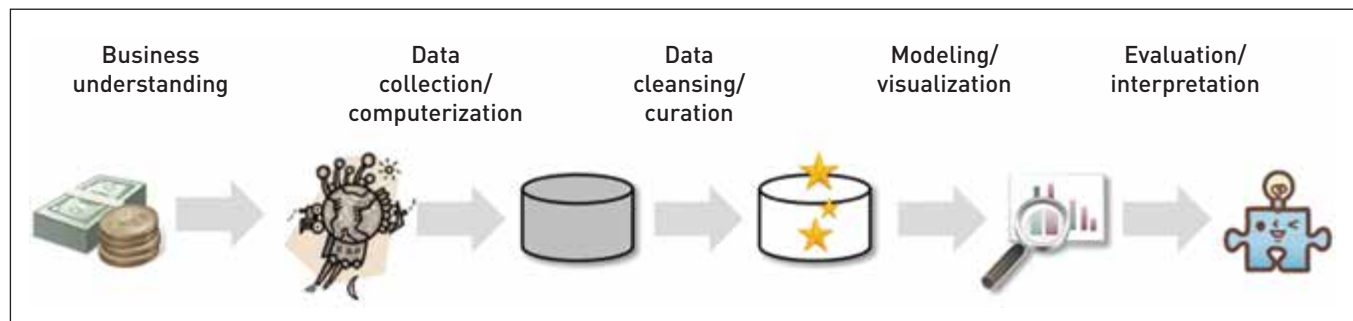


Fig. 1: Data Analysis Process

approach of omnipresent processing. The difficulties of huge information incorporate "its investigation, catch, information curation, look, sharing, stockpiling, exchange, perception, and data security". Enormous information is critical as it can enable us to accumulate, store, oversee, and control immense sums information at the correct speed, at the perfect time, to pick up the correct bits of knowledge. Big data stages enable us to virtualize and store information effectively. This is done most cost-adequately through cloud-based capacity.

Big data is commonly separated by three qualities:

- "Volume: How much information there is,
- Velocity: How quick that information is prepared,
- Variety: The different kinds of information".

2. Digital Context Of Crowdsourcing

Presently we look at the setting which made it feasible for publicly supporting to turn out to be such a vital marvel, the data society. In data society the creation, conveyance, use, reconciliation, and fundamentally even the treatment of data is a vital financial, political, social and social action. Data society can be estimated by a few markers: mechanical, financial, word related, spatial, and social pointers. Individuals who participate in (share of a sustenance, sedate, drink, and

so forth, which means taste, attempt expend) in this type of society are called advanced natives. Data society appears where broadband web, equipped for symmetrical associations is available both in physical and social spaces.

Another wonder identified with publicly supporting is Web 2.0. Web 2.0 administrations enable clients to collaborate and cooperate as makers of client created content utilizing ICT. Web 2.0 forms a virtual network of prosumers (buyers and makers), which are not restricted any more to being simply shoppers of substance. Instances of Web 2.0 incorporate long range "interpersonal communication destinations (Facebook), websites (Tumblr), smaller scale online journals (Twitter), wikis (Wikipedia), video sharing locales (YouTube), and among others, mashups (Google Maps)".

With crowdsourcing, the problem-solving task is outsourced to an undefined public (the crowd) through an open call via a Web-based business model. This approach has spawned a new breed of crowdsourcing technology platforms (e.g., Mechanical Turk, crowdSPRING), service providers (e.g., Kaggle, CrowdFlower, InnoCentive), and "crowdworkers" (individuals who participate in crowdsourcing initiatives for a living).

3. Crowdsourcing Big Data

Crowdsourcing is an imaginative methodology in the time of big data as

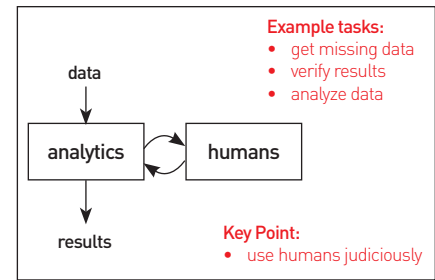


Figure 4: Crowdsourcing Big data Analytics

it improves appropriated handling and huge information examination.

Enormous advantages can be harvested by blending up publicly supporting with huge information:

1. Crowdsourcing big data enables associations to spare their interior assets - Why procure over qualified staff for huge information forms that publicly support workforce can handle all the more proficiently, rapidly and cost adequately.
2. Crowdsourcing big data enables associations to profit by the human component Content balance and assessment investigation from criticism of clients, social updates, surveys or remarks with publicly supported workforce results in exceedingly exact, significant and important bits of knowledge when contrasted with machines.
3. The appropriated idea of publicly supporting guarantees that enormous information is handled at an unforeseen speed which would not be conceivable to accomplish in-house.
4. Associations can fabricate applications dependent on constant examination as publicly supported workforce produce enormous information investigation at ongoing. Endeavors don't need to be made a fuss over being unfashionably late to the huge information party.

3.1 How crowdsourcing helps facilitate the procedure of big data analytics?

- Generally, an information researcher invests 78% of his energy in setting up the information for enormous information investigation. Therefore, a smart and financially savvy system



Figure 2: Real world examples of Crowdsourcing

for enormous information organizations is hand over the unstructured informational collections to a very much oversaw publicly supporting stage so the group will educate all the more concerning the data contained inside the information focuses gathered. For instance, before the examination the group can tell whether the information focuses are a Tweet or updates from Facebook and whether it conveys a negative, positive or impartial meaning.

- Crowd gives structure (archive altering, sound translation, picture comment) to enormous information in this manner helping experts improve their investigation prescient models by 25%.

- Crowdsourcing alongside enormous information examination can help uncover concealed bits of knowledge from scattered however associated data rapidly.
- Big information issues can be comprehended with more exactness with publicly supporting as a dependable medium.
- The results from the group can be utilized by information researchers to improve the productivity of the AI calculations.

3.2 Crowdsourcing context:

- Crowd — An individual or groups dealing with a movement and finishing it with zero ability to see to different people or groups
- Community — Individuals or groups dealing with a movement with some dimension of perceivability to

different people and groups

- Competition — Individuals or groups taking a shot at and finishing a movement autonomously (just a single victor)
- Collaboration — Individuals or groups taking a shot at parts of a movement and adding to its finish (everyone wins)

4. Conclusion

Crowdsourcing and Big data analytics together can enable associations to abuse data for settling on educated business choices that are a commendable journey. Crowdsourcing data is an effective way to seek the help of a large audience usually through the internet to gather information on how to solve the company's problems, generate new ideas and innovations. ■

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, 6 Technical articles in CSI Communications Magazine and over 100 publications in highly cited Journals and Conferences. His professional awards include: 100 Inspiring Authors of India, Deloitte Innovation Award, Cash Prize Rs.10,000/-, from Deloitte 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. Jayasudha Subburaj, dynamic and enthusiastic teaching professional with proven training abilities in the area of emerging technologies. She received Ph.D., degree in Computer Science and Engineering from Anna University, Chennai, in 2015 and M.E., in Computer Science and Engineering from Karpagam University, Coimbatore. She completed her undergraduate studies at Bharathiar University. She is currently working as a Placement officer at Sri Krishna College of Engineering and Technology, Coimbatore. She liaisons with the various industry personnel and persuading them to visit the college campus for conducting recruitment drives. She also provides career counseling for the students about the areas the students need be trained to make them industry-ready and employable. Her professional interests focus on software engineering, wireless networks and computing methods. In addition she serves as an Assistant Professor in the department of Master of Computer Applications. She was recently honored as Best Contributor award towards Students Career from Women In Product Group and GUVI.



Dr. J. P. Ananth is a Professor at Sri Krishna College of Engineering and Technology. He received B.E and M.E in Computer Science and Engineering from Manonmaniam Sundaranar University, and Doctor of Philosophy (Ph.D) in Computer Vision and Pattern Recognition from Sathyabama University, India. He has published research articles in highly cited journals and conferences. He has authored a book in Compiler Design. He is a Reviewer and Editorial Board Member of many International Journals. Also acted as Session Chair and Advisory Board Member in many Conferences. His research interests include Data Analytics and Computer Vision.



Machine Learning and Analytics

► Nasib S. Gill and D. Sehrawat

Department of Computer Science & Applications, Maharshi Dayanand University, Rohtak, Haryana (India). Email: nasibsgill@gmail.com

1. Introduction

Over the past several decades, digital computers have changed the perspective in almost every segment of the economy. Earlier it was a dream that one day machine would be able to learn but in modern life, almost all aspects are being changed in some way by big data and machine learning. For example, Netflix makes guess about one's movies interest and Google from search histories knows what people want to know [1], Facebook identifies faces in photos [2], our webpages are offering ads of our interest, emails detect spam, speech and text recognition. These are some of the persuasive examples of our daily life driven by the most dynamic and progressive machine learning. Today almost all of the industries are making benefits from machine learning applications including automobiles, genetics, health care, finance, banking, government, retail, education and etc. Machine learning helps these industries by automating procedures, reducing processing time, providing more accurate and quicker decisions.

Machine learning is a buzzword floating around nowadays. It is one of the most exciting subfields of research in computer science and engineering.

Machine Learning allows applications to predict more precise and accurate outcomes without explicitly programmed. It works by developing procedures that take input data and then by applying statistical analysis on the data, it predicts an output. Machine learning processes are similar to that of data mining. Both involve searching patterns through data and updating outputs according to new inputs.

2. Machine Learning

The term Machine Learning was coined in 1959 by Arthur Samuel, an American pioneer in the field of artificial intelligence and computer gaming.

Artificial Intelligence (AI) is a broad area of science which performs mimicking of human abilities. Machine

learning being a specific subset of AI is the field of science where computers automatically learn without being explicitly programmed. Machine learning provides the ability to learn from experiences and brings computers more similar to humans. It focuses on accessing data, making observations and adjusting actions accordingly. The data can be obtained from direct experiences, or instructions [3]. Machine learning analyzes this massive data and looks for patterns to make stronger decisions in the future. This process is known as building mathematical data from sample data/ training data to provide decisions or predictions without being explicitly programmed to perform tasks. Furthermore, it also provides solutions to those problems for which it is not feasible to develop an algorithm to perform the task like computer vision.

This is because of machine learning that faster and more accurate results are provided in order to categorize profitable opportunities or dangerous risks.

3. How does Machine Learning work?

Machine Learning algorithms are often categorized as supervised, unsupervised and reinforcement learning [4].

1. Supervised machine learning algorithms work by receiving input data, desired output, and feedback from data scientist/ data analyst and update the algorithms to improve the precision of predictions. Data scientists, during algorithm training, make decisions about which variables or features should be used and analyzed to improve the predictions. The algorithm after training will apply to new data for prediction.
2. Unsupervised machine learning algorithms require no training. Instead, they use deep learning, an iterative approach for reviewing data and reaching conclusions.

Furthermore, unsupervised learning algorithms using neural networks automatically find correlations between many variables and can solve even more complex processing tasks than that of supervised learning systems.

3. Like unsupervised learning, reinforcement learning is not provided with input/output pairs but like supervised learning, feedbacks in the form of rewards or punishments are provided in reinforcement learning.

Steps in Machine Learning process:

1. Identify relevant data sets and prepares them for analysis.
2. Choose an appropriate machine learning algorithm to use.
3. Build an analytical model on the chosen algorithm.
4. Training of model on input data sets.
5. Executing model to generate scores and results.

Requirements of good machine learning system

- Abilities to prepare data
- Advanced algorithms design
- Automation and iterative processes
- Scalability
- Ensemble modeling

4. Relation to Data Mining

With big data era, computers and internet applications have spawned huge high dimension data at high speed like photo, video, text, and voice attained from the ever-increasing Internet of Things and cloud computing. This has posed a challenge for data analysis and decision-making processes [5]. With this big data, interests in machine language have spurred.

Machine learning and data mining overlap significantly, but while data mining finds previously unknown patterns in the data, machine learning emphasizes on the prediction by learning from the training data. Data mining make use of several machine learning methods; whereas, machine

learning employs data mining as a preprocessing step for improved accuracy [6].

Within machine learning, data mining is a field of study which emphasizes on exploratory data analysis through unsupervised learning. Within the field of data analytics, machine learning provides algorithms for complex models for prediction, an area known as predictive analytics. These analytical models allow engineers, analysts, researchers, and data scientists, to "produce reliable, repeatable decisions and results" and uncover "hidden insights". This is all done by identifying relationships and patterns/ trends in the historical data set (input data).

5. Impact of Machine Learning on Everyday Life

Machine learning is adequately young and is still rapidly growing with several underexplored research opportunities [7]. Today's machine learning is not like machine learning of the past. In the past decade, machine learning has given us self-driving cars, practical speech recognition, and effective web search but today, it has gained fresh momentum and is so pervasive that we are actively using it many times a day even without knowing it. This is because of innovative computing technologies that complex mathematical calculations can be applied automatically to big data more quickly and repeatedly.

Few broadly publicized examples of

machine learning applications include neuroscience, Internet of Things, online shopping, online recommendation, served ads, photo tagging, personalized marketing, fraud detection, health care (diagnose patients), predictive maintenance, email filtering, spam filtering, computer vision and building news feeds, brain-machine interfaces, intelligent game playing, natural language processing, study of biological networks, network security threat detection, genome sequencing data sets, personalized digital media, streamlined logistics and distribution, security and privacy, robot locomotion, and etc. Besides, machine learning is also powering virtual assistant technology by combining several deep learning models to provide pertinent context and interpreting natural speech.

Machine learning is helping us to live happier, healthier and more productive lives. Machine learning is more valuable when one wants to predict future events when people know what they want but don't know how to take the decision and what are the necessary input variables.

6. Conclusion

The social and ethical impact of machine learning will continue to stir the world's dreams. Its ability to work on big data and predicting outputs for complex problems would help in creating a new framework for problem-solving. It is clear from the above discussion that machine learning along

with data mining is capable of solving a wide range of problems. Machine learning is a rapidly growing technology with numerous underexplored research opportunities.

7. References:

- [1] Beam, A.L. and Kohane, I.S., 2018. Big data and machine learning in health care. *Jama*, 319(13), pp.1317-1318. [10.1001/jama.2017.18391](https://doi.org/10.1001/jama.2017.18391)
- [2] Mullainathan, S. and Spiess, J., 2017. Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2), pp.87-106. [10.1257/jep.31.2.87](https://doi.org/10.1257/jep.31.2.87)
- [3] Erickson, B.J., Korfiatis, P., Akkus, Z. and Kline, T.L., 2017. Machine learning for medical imaging. *Radiographics*, 37(2), pp.505-515. <https://doi.org/10.1148/rgr.2017160130>
- [4] Zhou, L., Pan, S., Wang, J. and Vasilakos, A.V., 2017. Machine learning on big data: Opportunities and challenges. *Neurocomputing*, 237, pp.350-361. <https://doi.org/10.1016/j.neucom.2017.01.026>
- [5] Cai, J., Luo, J., Wang, S. and Yang, S., 2018. Feature selection in machine learning: A new perspective. *Neurocomputing*, 300, pp.70-79. <https://doi.org/10.1016/j.neucom.2017.11.077> [10.1126/science.aaa8415](https://doi.org/10.1126/science.aaa8415)
- [6] L'heureux, A., Grolinger, K., Elyamany, H.F. and Capretz, M.A., 2017. Machine learning with big data: Challenges and approaches. *IEEE Access*, 5, pp.7776-7797. [10.1109/ACCESS.2017.2696365](https://doi.org/10.1109/ACCESS.2017.2696365)
- [7] Jordan, M.I. and Mitchell, T.M., 2015. Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), pp.255-260. [10.1126/science.aaa8415](https://doi.org/10.1126/science.aaa8415)

About the Authors



Dr. Nasib Singh Gill is at present senior most Professor of Department of Computer Science & Applications, M. D. University, Rohtak, India and is working in the Department since 1990. He earned his Doctorate in Computer Science in the year 1996 and carried out his Post-Doctoral research at Brunel University, West London during 2001-2002. He is a recipient of Commonwealth Fellowship Award of British Government for the Year 2001. Besides, he also has earned his MBA degree. He has published more than 245 research papers in reputed National & International Journals, Conference Proceedings, Bulletins, Edited Books, and Newspapers. He has authored seven books. He is a Senior Member of IACSIT as well as a fellow of several professional bodies including IETE and CSI. He has been serving as Editorial Board Member, Guest Editor, Reviewer of International/National Journals and a Member of Technical Committee of several International/National Conferences. He has guided so far 9 Ph.D. scholars as well as guiding about 7 more scholars presently in the areas – IoT, Machine Learning, Information and Network Security, Computer Networks, Measurement of Component-based Systems, Complexity of Software Systems, Decision Trees, Component-based Testing, Data mining & Data warehousing, and NLP.



Ms. Deepti Sehrawat has passed Master of Computer Science and Applications from Department of Computer Science & Applications, Maharshi Dayanand University, Rohtak, India in 2008 and Master of Philosophy from C.D.L. University, Sirsa in 2009. She is currently pursuing Ph.D under the supervision of renowned academician and researcher – Professor Nasib Singh Gill of M. D. University. She has published more than 25 research papers in reputed National and International Journals and Conference Proceedings including IEEE. Her main research work focuses on IoT, Lightweight Cryptography Algorithms, Network Security and Privacy, Big Data Analytics and Data Mining. She has 8 years of teaching experience.

Role of Data Analytics in Health Care Sector – A Research View

► T. Ramesh

Assistant Professor, Department of CSE, School of Engineering, Presidency University, Karnataka. Email: ramesh.t@presidencyuniversity.in

► V. Santhi

Associate Professor, School of Computing Science & Engg. VIT Univ., Vellore-632104, Tamilnadu. Email: vsanthinathan@gmail.com

Data Analytics – Introduction

Data analytics (DA) is the process of inspecting data sets in order to get the inference by which conclusions can be drawn by making use of specialized systems and software. Especially data Analytics are playing huge role in making business decisions which in turn helps industries to increase their turn over. It has huge market around the competitive world which can help in providing better services for the customers and it can be applied effectively in all the domains such as healthcare, banking, stock markets, Transportation, Banking, Logistics and supply chain management etc.

Role of Data Analytics in Health Sector

Data Analytics plays a vital role in the field of healthcare sector, where lot of current challenges will get resolved using this technology. Some challenges in healthcare sector are such as Reducing of Medical Errors in the diagnosis of diseases identification, improving the patient safety, increasing the revenue of hospitals etc.

In this article we focus on Reducing Medical Errors in hospitals as the expenditure for treatment is gradually increasing in an exponential manner. This becomes a huge burden for lower middle class people as they are not in a position to bear the expenses incurred towards a treatment. Before data analytics comes in to the picture, there is lack of clarity in diagnosing various types of diseases, sometimes it may lead to failure. So, data analytics with the combination big data plays a vital role in healthcare analytics.

Why is it important to reduce medical errors

The reduction of medical errors is very important for any hospital in order the reputation & sometimes these medical errors lead to a patient death.

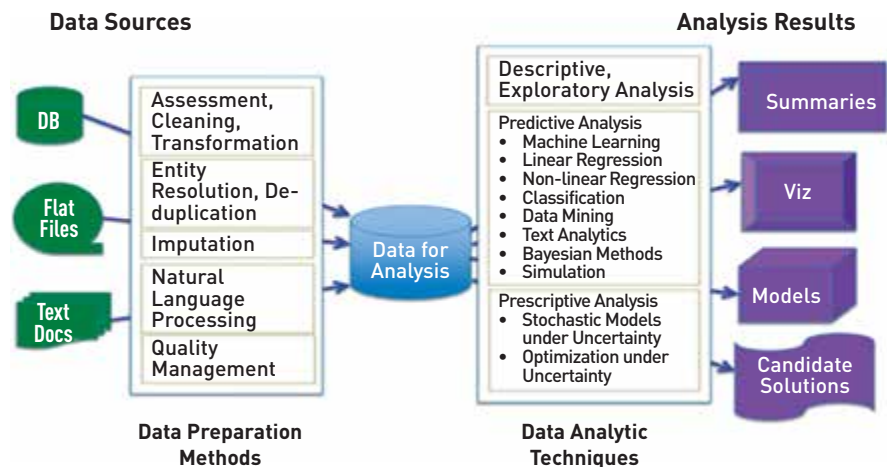


Fig. 1: The evolution from the data sources to analysis results passes through several steps

As Chances of occurring medical errors at various stages starts from admission to discharge from the hospital, there is a huge requirement for data analytics.

Need for reduction of errors in healthcare- At Present

At present there is an enormous amount of challenges exist in healthcare as the number of diseases are getting increased and still for particular diseases there is no medicine available outside, even though if it is available poor people are not in position to afford those medicines. So, the data analytics will be very much useful in healthcare in terms of prediction of diseases at the early stage will help them to reduce expenses for treatment towards a particular disease. Data Analytics will help not only in predictive analysis and also in increasing operational efficiency etc. Medical errors in general costs very much high as it may cost patients death.

Benefits of Error Reduction in healthcare over organizations:

Error reduction in healthcare will help organizations to provide better services to all the stakeholders including

patients, doctors, administrative people etc. In the process of applying analytics, various types of data need to be analysed including structured and unstructured data by which amount of precision will increase for better results. Upon accessing data from various sources such as data bases, text documents etc. which need to be under gone various pre-processing techniques such as cleaning, transformation etc. before applying analytics over it. So, data analytics will help organizations for both prescriptive and descriptive analysis as mentioned in fig 1. where it shown evolution of data sources by passing through various steps in which analytics can be applied effectively.

The outcome of the data analytics in healthcare is in the form of summary reports or any other models such that one can easily understand the inference from huge volume of data. So, there are so many methods to apply data analytics over healthcare.

Conclusion:

The reduction in medical errors in health care sector will help organizations to increase their revenue

such that they can provide better services for all the stakeholders. So, the data analytics will provide a solution for all the challenges in healthcare such that amount of errors can be minimized gradually. Still so much research is

going on in healthcare where data analytics will provide a solution for upcoming challenges.

References:

[1] <https://searchdatamanagement.techtarget.com/definition/data-analytics>

- [2] <https://www.digitalvidya.com/blog/data-analytics-applications/>
 [3] <https://www.brookings.edu/research/the-opportunities-and-challenges-of-data-analytics-in-health-care/>
 [4] <https://images.app.goo.gl/E51Yq8maFoZeXw4Q6>

About the Authors



Mr. T Ramesh currently working as an Assistant Professor in Department of Computer Science & Engineering at Presidency University, Bangalore Karnataka. He has completed B. Tech, M.E, in JNTU University Hyderabad and Bangalore University, Karnataka in 2005 and 2012 respectively. He is currently doing his research in health care sector using big data analytics under VIT University, Vellore. He published many papers in many reputed journals. His research area includes Big Data analytics, Data Mining and Machine Learning.



Dr. V Santhi currently working as an Associate Professor and also serving as a HOD for B.Tech Programme in the School of Computing Science & Engineering at VIT University, Vellore. She has completed her Ph.D in the field of Image Processing at VIT University, and she has published so many papers in reputed journals across the globe. Her research interests are Image Processing, Steganography, Digital Water Marking and Data Mining.



CSI PUBLICATIONS TEAM



DR. D. D. SARMA

Chairman
Publications Committee

2-16-104. Prasanthnagar, Uppal
Hyderanad 500039

☎ 096180 24570 / 77021 18446

@ drddsarma@gmail.com

CSI COMMUNICATIONS



PROF. (DR.) S. S. AGRAWAL

Chief Editor

☎ 98101 98642

@ ss_agrawal@hotmail.com



DR. RITIKA WASON

Editor

Associate Professor, Bharati Vidyapeeth's,
Institute of Computer Applications and
Management (BVICAM), A-4, Paschim Vihar,
New Delhi-63.

☎ 9818411596

@ rit_2282@yahoo.co.in

CSI JOURNAL OF COMPUTING



DR. R. R. DESHMUKH

Chief Editor

Van-Lakxmi, D-8, Shevanti, Tirupati
Park, N-4, Cidco, Cidco Colony
(Aurangabad) Aurangabad-431003

☎ 9423147466

@ rrdeshmukh.csit@bamu.ac.in



DR. BROJO KISHORE MISHRA

Editor

Professor, Dept. of CSE, School of Engg. &
Tech., GIET University, Gunupur-765022,
Odisha

☎ 9437875808

@ bkmishra@giuet.edu
brojomishra@gmail.com



DR. VISHAL JAIN

Editor, CSI Adhyayan

Associate Professor, Bharati Vidyapeeth's,
Institute of Computer Applications and
Management (BVICAM), A-4, Paschim
Vihar, New Delhi-63.

☎ 98999 97263

@ drvishaljain83@gmail.com

Evolution of Data Analytics in the IOT Era...

► S. Shenbagavadivu

A.P. (Sr.G) in the Dept. of IT at SRM Valluammai Engineering College

► M. Senthil Kumar

Associate Professor in CS and Engg.
Department at SRM Valluammai Engg.
College of Tamil Nadu

► B. Chidambara Rajan

Professor/Principal at SRM Valluammai Engineering College, Affiliated to Anna University, Chennai

It is a well-known fact that the Technological Advancement in the IOT is reaching greater heights each and every day. In the 1990's about 1 billion devices were connected to the internet. With the development of mobile wave, 2 billion devices were connected to the IOT in the year 2000. More than 6 billion devices have been connected to the IOT in the past two decades. As per the estimation from Cisco the data generated from IOT would reach 600ZB per year by 2020. The growing market trend has generated greater demands in the field. IOT collects data from various sources such as sensors, smart devices, and micro controllers and on board units etc., now it would be a big challenge for us to collect, store and handle those data. Moreover it is very important for us to convert the data into useful information that can be made available to the end user. This is achieved by proper data analytics techniques.

Introduction

We can make the information available to the end user only after properly collecting and analyzing the data. Based on the organization's need and the type of data that is collected we need to select the appropriate data analytics techniques. Data analytics can be defined as the amount of data that the IOT device receives to understand from the scratch, to analyze about the working process, to aid the organization to take decision in the perfect manner. Every day 2.5 quintillion bytes of data generated, only 0.5% is analyzed and used. There are fresh opportunities available for the field of data science.

Data Analytics Techniques

There are four common analytics .Following figure shows the techniques.

1. Descriptive Analytics

Descriptive Analytics is used to provide answers to the questions for what has happened so far. Used to express about the pros and cons of current processing methods. It is used to evaluate the organization performance and provides the information in an effective manner.

2. Diagnostic Analytics

Descriptive Analytics is used to analyze historical data for answering the question of "Why something happened?".It is discriminated by the techniques such as data discovery, drill down, data mining and correlations.

3. Predictive Analytics

Predictive Analytics is used to make predictions about future events. More micro level predictions can be done. For example predicting all the items that the customer will purchase together.

4. Prescriptive Analytics

The main goal of prescriptive Analytics is to decide the actions needed to undergo to eliminate the problems that will happen in future. It recommends number of possibilities for solving the problems that will happen in future. This technique can be applied on data from various data sources such as real time data, historical data, big data and transactional data. It is very useful to deliver the right product at right time to satisfy the customers.



Fig. 1: Data Analytics Techniques

Process involved in Data Analytics

Step 1 (Raw Data): Used to determine the requirements of data and grouping of data. Generally data can be grouped together based on the sex, age and income.

Step 2 (Data Collection): Describes the process of collecting data. Data collected from various sources such as Surveys, cohort method, case control method, ecological method and by conducting experiments. Using appropriate collection method we will perform analytics

Step 3 (Data Organization): Once the data collected it must be organized so that it can be analyzed in an effective manner. Data can be organized using the following methods:

- Location
- Alphabet
- Hierarchy
- Time
- Category

Step4 (Data Cleansing): Before analyzing the data we need to clean up the data. It will clear any errors in data before it goes on to data analytics phase. It also ensures that there is no duplication, and if any state is incomplete.



Fig. 2: Process involved in Data Analytics

Data Analytics Tool

R Programming Language

R is leading analytics tool used in the industry especially for statistical and data modeling to perform predictive analytics. For example "what item would be purchased by the customer after 1 Year?". It includes techniques such as linear and non linear modeling, classification techniques, clustering, graphical techniques and others. It can easily manipulate the data and represent it in an effective manner. Google and Face Book use R programming to know about the analysis on status updates.

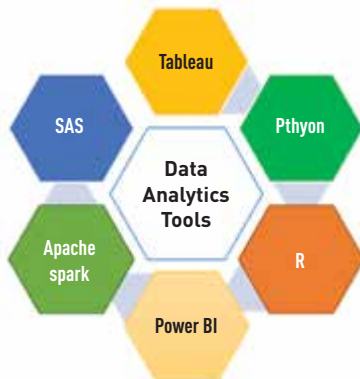


Figure 3: Data Analytics Tool

Tableau Public

It is a free visualization tool and most widely used in Business Intelligence industry. It helps to represent raw data in an understandable format .It can extract simple data base such as PDF. Excel to a complex databases such as

AWS, Microsoft Azure, Google Cloud SQL in an effective manner..

Python

Python is easy to learn and it is a powerful programming language. It is as that of PHP, Ruby and Java Scripting. It can run on different platforms such as Mac, Linux, windows and ResberryPi. It is used to create WebAPP on server side, workflows and to handle big data.

SAS (Statistical Analysis System)

It is used to perform advanced analytics and it can predict behavior and manage data from various sources. Social media and market analytics can be widely for profiling customer and prospects.

Apache Spark

Apache Spark is used to perform large scale data processing. It is used to write applications in SQL Shell, R, Scala, Python and Java. It can run on Apache Mesos, Kubernetes, and Hadoop, stand alone or cloud.

Power BI

It is a Business Analytics solution provided by Microsoft. It helps the organization to gather data from multiple sources in to one dashboard and thereby it identifies the problems. We can get real time alerts with PowerBI mobileApp and can make our operation more efficient.

Conclusion

Data Analytics helps organizations with actionable intelligence via industry-leading technology, analytics and

sophisticated enterprise intelligence applications.

By applying appropriate analytics techniques as mentioned above revenue of individual or organization is enhanced which will in turn act as a major factor of our country's growth. This also allows executives and operators alike to navigate via a single pane of glass, in near real-time, giving everyone easy access to metrics that optimize inventory, reduce waste, and shorten order-to-cash and time-to-ship metrics. The quality of work done will be as per expected standards by using advanced analytics tools. Risk will be predicted at the earlier stages itself and hence the failure of a system will be avoided. Design blames production, accounting blames sales. Eliminate the cancer of the blame-game by giving everyone a shared picture of progress towards the goals. Empirically identifying constraints drastically reduces finger pointing and empowers organizations to make critical decisions on the ground.

References:

- [1] www.slideserve.com/syscraftinc/impac-of-data-science
- [2] www.sisense.com/
- [3] imarticus.org/
- [4] www.scnsoft.com/blog/4-types-of-data-analytics
- [5] www.investopedia.com/terms/d/data-analytics.asp
- [6] www.albert.io/blog/data-collection-methods-statics
- [7] www.analyticstraining.com

About the Authors



S. Shenbagavadivu [CSI Membership No: I1504751] working as an A.P(Sr.G) in the Department of Information Technology at SRM Vallaimmi Engineering College. She has 12 years of teaching experience and her area of research includes IOE, IOT, Big data and Data Analytics.



Dr. M. Senthil Kumar (LM- I1504760) is currently working as an Associate Professor in Computer Science and Engineering department at SRM Valliammai Engineering College of Tamil Nadu. He is a CSI-Student Branch Counselor of the College. His research interests are in IOT, Big Data, Software Engineering and development of new tools for effort estimation.



Dr. B. Chidambara Rajan (LM-00063930) working as a professor/Principal at Valliammai Engineering College, Affiliated to Anna University, Chennai. He has 20 years of teaching experience in government and reputed private institutions. He is a member of professional societies like CSI, IEEE, IETE, IEI, ISTE, ISOI, etc. He has published several technical papers in national and international journals and conferences His research interests include IOT,Big Data, Software Engineering and Networking.

CSI EXECUTIVE COMMITTEE

2019-20/21

EXECUTIVE COMMITTEE



President

Prof. A. K. Nayak

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



Vice President (President Elect) (2019-20)

Mr. Ram Krishna Vyas

70, Sanskrit Nagar Society, Plot No-3
Sector -14, Rohini, Delhi -110085
(T) 011-2786 6259
(M) 98105 92760
(E) rkvyas_99@yahoo.com



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



Hony. Treasurer (2019-21)

Dr. Durgesh Mishra

H-123-B, Vigyan Nagar,
Annapurna Road, Indore-452009
(M) 98260 47547
(E) drdurgeshmishra@gmail.com



(Immediate Past President-2019-2020)- Continued

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 (2019-21)

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 (2019-21)

Md. Shams Raza

Ignou Programme Centre
St. Xavier School Campus
Gandhi Maidan, Patna - 800 001
(M) 94308 28918
(E) s_raza2000@yahoo.com



Region-III (2017-19)

Mr. Jayant S. Binde

48 Anupam Nagar,
Near Jiwaji University,
Gwalior 474011
(M) 94251 09013
(E) jayantbhide_2000@yahoo.com



Region-IV (2018-20)

Er. Nachindra K. Behera

Plot No. 223, Prachi Enclave-II
Chandrasekharpur
Bhubaneswar-751016, Odisha
(M) 99373 75673 / 94388 38527
(E) nachindrabehera@gmail.com

CSI EXECUTIVE COMMITTEE

2019-20/21

REGIONAL VICE PRESIDENT



Region-V (2019-21)

Prof. M. Surendra Prasad Babu

Retd. Professor Dept. of Computers
Science, Systems Engg., Andhra
University, A U Engg College
Visakhapatnam-530003
(M) 98480 56369
(E) profmspbabu@gmail.com



Region-VI (2018-20)

Mr. Pradeep Rathi

E-401- Arshie Complex
Panch Marg Versova
Mumbai-400061.
(M) 98202 93998 / 98200 19503
(E) pjrathi61@gmail.com



Region-VII (2019-21)

Dr. M. Sundaresan

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



Region-VIII (2019-21)

Mr. Konatham Mohan Raidu

Director, Informatics India
1-9-1113/1/A/A, Vidyanagar,
Hyderabad 500044
(M) 9985033566
(E) rvp8@csi-india.org
kmraidu@informaticsindia.com

DIVISION CHAIRPERSONS



Division-I (2019-21)

Mr. Manas Ranjan Pattnaik

Plot No. N-24, 25 Chandaka Industrial
Estate, Patia, KIIT, Bhubaneswar
(M) 07873099999
(E) manas@antheimgt.com



Division-II (2019-21)

Col. Balraj Anand

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



Division-III (2019-21)

Prof. Suresh Chand Tyagi

13/164, Vasundhara, Kaushambi
Ghaziabad-201012
(M) 98110 54645
(E) sct_35_2000@yahoo.com
scttyagi1963@gmail.com



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. Subhash Chandra Yadav

Associate Professor and Head Dept. of CSI & Tech.,
Central University of Jharkhand, Ratu-Lohardaga Road,
Brambe, Ranchi-835205
(M) 94158 70233
(E) scy123@gmail.com

NOMINATIONS COMMITTEE (2019-2021)



Mr. Arvind Mohan Nayak

203 Napier Grand,
Near Home Science College,
Napier Town Jabalpur-482002
(M) 94280 07518
(E) amnayak@hotmail.com



Mr. Deepak Sharma

House No. D-21,
Nirala Nagar,
L. P. Marg, Lucknow-226020
(M) 9795121212
(E) deepak.sharma@nic.in



Dr. Ratnadeep Deshmukh

Van-Lakxmi, D-8, Shevanti, Tirupati Park,
N-4, Cidco, Cidco Colony (Aurangabad)
Aurangabad-431003
(M) 9423147466
(E) ratnadeep_deshmukh@yahoo.co.in
rrdeshmukh.csit@bamu.ac.in



Summer Offer for Life Membership

I, hereby, apply for new membership. On approval of Membership, I shall abide by the Constitution & Byelaws of the Society and the Code of Ethics. Please send a good quality minimum 300 x 300 pixels / passport size photograph to swapnil@csi-india.org to be used for making your CSI Membership Card (photo required only for Life Members) along with a copy of Voter ID / Aadhar Card / PAN Card / Driving Licence

I. Select the membership type

Indian ☐ International ☐

Please tick for Membership period

One Year ☐ Two Years ☐ Three Years ☐ Four Years ☐ Life ☐

Paste your (recent)
one passport size
photograph here.

II. Personal information:

Please fill in your personal information so that we can serve you better

Title of the applicant Mr. ☐ Miss ☐ Mrs. ☐ Dr. ☐ Prof. ☐

First Name Middle Name Last Name

Name you would like to be printed on CSI ID card

Date of Birth

Gender

Primary Email ID

Secondary Email ID

Phone No. (Residence)

STD Code

Phone

Mobile (Mandatory for domestic membership)

Highest Academic Qualification:

Year of Passing:

University/Institute:

Total work experience in years:

Current Employment & Designation:

CSI Communications: Hard Copy ☐ Soft copy ☐
(Please tick)



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 – II
Version 2.4
w.e.f. 01st July, 2017

III. Mailing address (BLOCK LETTERS):

Address line 1

Address line 2

Address line 3

Pin Code

City

State

Country

(City, State and Country to be filled in only for international address)

Name of the Chapter to be attached

IV. Payment details:

Please specify Mode of Payment: NEFT/ Demand Draft/Cheque]

If payment made through NEFT – Transaction details Date

₹

UTR No.

If payment made through Demand Draft DD / Cheque payable at par at Mumbai should be drawn in favour of “Computer Society of India”.

Cheque ☐ DD ☐ Cash ☐ (Please tick as applicable)

Amount Paid in ₹ / \$

Cheque / DD No.

Dated

Drawn on Bank Name

Branch Name

Please fill following details if it is deposit in Axis Bank.

Date of Deposit

Mode of Deposit

(Please tick as applicable)

Axis Deposit branch name

Bank Details

A/c Name: Computer Society of India

Bank Name: Axis Bank Ltd.

A/c type: Saving

A/c No: 060010100082439

IFSC code: UTIB0000060

Bank Address: Aman Chambers Ground Floor, Opposite New Passport Office, Veer Savarkar Marg Worli, Mumbai 400 025

GSTIN- 27AAATC1710F1ZC

SAC-9995



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 – II

Version 2.3

w.e.f. 01st July. 2017

Attach photocopy of Pay-in-slip with application form and write your Name, Contact no. on the reverse side of the Cheque / DD / Pay-in- Slip.

III. Mailing address (BLOCK LETTERS):

I affirm that as a CSI member, I shall abide by the Code of Ethics of the Computer Society of India (CSI). I, further, undertake that I shall uphold the fair name of the Computer Society of India by maintaining high standards of integrity and professionalism. I was not a member of CSI earlier. I am aware that my breach of the Code of Ethics may lead to disciplinary action against me under the Byelaws and rules of the CSI. I, hereby, confirm that I shall be bound by any decision taken by the CSI in such matters. Further, 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: / /

Place:

Signature : _____

FOR OFFICE USE ONLY

Application received date : _____

Received by : _____

Application processed by : _____

Membership No.

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

VI. Membership Subscription Fees:

Life Membership Fee (irrespective of the age of the applicant)

(The membership Period is on Rolling Year basis)

CSI SUMMER OFFER Valid from 15 th May 2019 - 15 th July 2019	15% Discount on ₹ 11,800/- (including GST) = ₹ 10,030/-
Conventional Life Membership Fees (in instalments)	₹ 10,000 + 18% GST = ₹ 11,800.00 Can be paid in 4 equal instalments spread over 4 years*. The four cheques should be submitted alongwith the application form :- Each year ₹ 2,500.00 + 18% GST = ₹ 2,950.00 *Note: i. Three PDCs of the amount ₹ 2,950.00 are to be given in the first year itself, along with the Membership Application Form. ii. Membership shall be terminated with immediate effect, if the PDCs are not realized. iii. Additional liability, on account of any subsequent changes in the GST rule will need to be paid by the member.
Outside India	USD \$ 650

Note - Please attached below documents along with form.

1. Graduation Certificate
2. Highest Qualification Certificate

Monetize Dormant Data: 3 Steps for C-Suite Executives to Integrate Analytics into Decision Making

► Binesh Nair

Senior Consultant in Analytics and Information Management in Deloitte, Ireland. Email: binesh.nair3@gmail.com

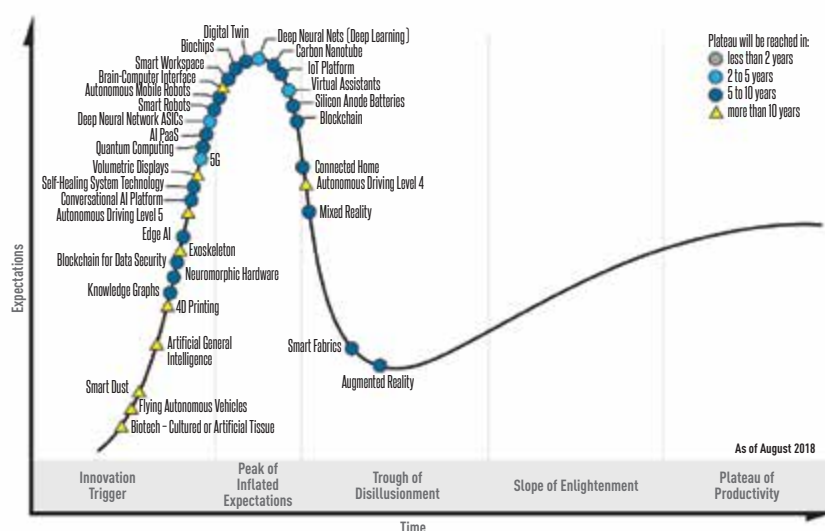
Although data analytics seem to be ubiquitous; the reality is that many organizations are still grappling to find value from the data lying dormant within them. A 3 - pronged approach presented in this article-creating SMART business use cases, building a well-balanced analytics team and establishing an insight-driven culture; can help organizations to capitalize on the insights generated from underlying data to drive business growth.

Organizations are often clueless when it comes to making sense of the endless list of “next big thing” technologies that seem to appear every other day in the industry. These technologies are often eulogized in conferences by key note speakers and trumpeted in industry-recognized publications. At the same time, it is amusing when each of these technologies is said to ‘revolutionize’ or ‘disrupt’ the industry! While many of these technologies may have the potential to revolutionize the industry over the next decade or so; unavailability of any reasonable business use cases makes it strenuous for organizations to estimate the impact these technologies will have on the industry in general and on the organization in particular. Gartner has brilliantly identified and visualized the stages of the emerging technologies in the Hype Cycle 2018 [see Fig. 1].

Gartner had dropped data analytics from its hype cycle for emerging technologies in 2015. Although this can be inferred as analytics as a technology being ubiquitous in 2019; the reality is that many organizations are still grappling to find value from the data lying dormant within them.

If your organization is still waiting to adopt analytics and is clueless about how to go about it; then you can take the following three steps to reap the full potential of your most valuable and at the same time grossly under-utilized

Hype Cycle for Emerging Technologies, 2018



Source: Gartner (August 2018)
© 2018 Gartner, Inc. and/or its affiliates. All rights reserved.

Gartner®

Gartner

Fig. 1: Hype Cycle for Emerging Technologies

asset: “data”. This 3-pronged approach can act as a roadmap for integrating analytics into your business.

Creating SMART Business Use Cases

Many organizations are often forced to implement emerging technologies in order to overcome the fear of being disrupted by the technology or being outplayed by the competitors! However, most of the times the stakeholders in such organizations either do not have a business case in hand or have only a

vague understanding of how emerging technologies can solve a business issue in hand.

The obvious challenge in implementing any ‘half-backed’ technology is that, it becomes too difficult to measure the impact of the technology and hence, becomes even more unreasonable to communicate the ROI on this technology!

Hence, in order to avoid costly mistakes, not just in terms of costs

but also, in terms of intangibles like, employee engagement, brand reputation or customer experience; it is paramount not only to identify a business use case but also, ensure that it adheres to the SMART framework. In other words, a business use case must be Specific, Measurable, Actionable, Relevant and Time-bound.

While data analytics appear to have been democratized with the availability of tools, talent and reproducible business use-cases; many organizations still seems to be clueless on how to gain actionable insights from data. For instance, more than 70 per cent of the organizations which participated in the 2018 Deloitte Insight Driven Organization (IDO) survey have challenges in generating value from data; with most of the efforts being ad-hoc and in silos. This is substantiated by the fact the most of the organizations are struggling to manage data; 66 percent of the respondents in the IDO survey identify data management as a key barrier to the adoption of analytics.

How can organizations create a use case? In order to answer this question, an organization has to identify an issue which is challenging the business growth. A business issue can either be a 'visible' business problem that is affecting the business like, consistent drop in sales or consistent churn of customers etc., or it can be an 'invisible' problem that is not directly affecting the business currently but, is rather stopping the business from delivering on its full potential. An instance of an 'invisible' problem is a scenario wherein the customer support team is working overtime to address customer's issues. While on the surface it may appear that the issue is a poor customer support team and hence, an impulsive solution will be to develop an innovative model to reduce the time taken to address a customer issue. However, an analytical approach is to do a 'root-cause' analysis to answer questions like, why there are many issues with the product in the first place? Are these issues redundant? Are there any hidden patterns?

Once a business issue has been identified and well-defined, it is important to identify if analytics can be used to address the issue? For instance, it may even be an issue with

the business process rather than lack of insights! If the solution requires data to be crunched then the first step will be to define objectives that can be monitored once an analytics solution has been implemented. In other words, it is critical to explain what deliverables will look like, preferably in business parlance. This essentially will help in communicating to a wider business audience without the fear of being lost in translation.

While any organization will have multiple key issues at any point of time, a best practice is to identify a pilot project from a set of potential projects based on cost, feasibility, time and resource restrictions against the potential business impact the project can make; as shown in Fig. 2. One of the benefits of this visualization is that, it allows the organization to prioritize the issue based on various parameters. For instance, if the purpose is to taking majority of the stakeholders into confidence, then it may be ideal to choose a project which can make a bigger impact. Similarly, if cost is a concern then the organization can choose smaller projects which can deliver quick wins.

A pilot project is also an effective way to gain the stakeholders' confidence in analytics and it also provides them enough teeth to buy-in other stakeholders or to communicate the potential impact to a wider organization.

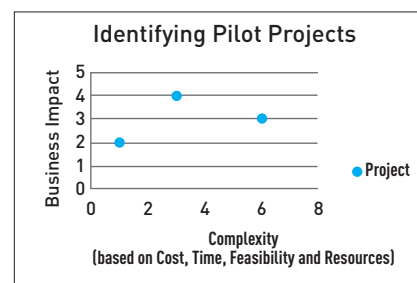


Fig. 2: Identifying Pilot Projects for demonstrating Impact of Analytics

Building the Right Team

Once the business issue has been identified and it has been translated into an analytics use case with well-defined deliverables; it is time for the organization to identify and define the required capabilities. Three broader capabilities will be required by any organization to integrate analytics into decision making: capability to plan and

execute the analytics strategy; capability to translate the business requirements into analytics problems and capability to communicate analytical findings as business insights. Some of these capabilities will be hard to find, based on the sector and the market in which the organization operates. While supply is definitely a concern for these multi-disciplinary skillsets; organizations also have to align their recruitment strategy and process to the analytical needs of the organization. For example, only 31 percent of the organizations in the 2018 Deloitte IDO survey have broken their analytical needs at a capability level and proactively embedded this into their hiring and talent development strategies and processes.

Capabilities can be present either within an organization or need to be acquired from the market. If capabilities are present in-house; then the organization will have to identify the level of training that will be required in order for the resources to start delivering on the project. Ideally, in a context wherein the organization itself is testing water in terms of integrating insights from an analytics solution into decision making; it is advisable to have atleast one senior leader who has both business know-how and technical knowledge in analytics. He will be the face of analytics in the organization and will play a key role as a translator between business stakeholders and technical analytics resources. Additionally, it is a best practice to have a data evangelist; typically a senior leader from the business side - who understands analytics, appreciates the impact that analytics can make in terms of delivering business outcomes and can manage the various stakeholders in the business.

If the capabilities is being searched outside, it is important to understand the kind of skills required in the analytics team. For instance, it is important to understand the distinction in the roles of a data engineer, a data scientist and a data analyst. While tool knowledge may seem to be important for a data engineer; attributes like critical thinking, business sense, ability to translate business requirements, telling stories using insights will be more important for a data analyst.

Organizations should restrict themselves from defining generic job descriptions without understanding the key capabilities required for that position.

Analytics is not about technology or tools; it is about being able to translate the business requirement into an analytics problem and being able to communicate the insights to business and also being able to convince the stakeholders to act on these insights.

Setting up an Insight-Driven Culture

Culture within an organization is defined by its people; which in turn is decided by the organization itself. One of the ways to identify priorities within an organization is to identify the influence of people who are in-charge of these priorities and the visibility of their business functions. For instance, high priority business functions will have influential people and high visibility. As per the 2018 Deloitte IDO survey, only 15 percent of the respondents have a Chief Analytics Officer or a Chief Data Officer. Only 37 percent of the organizations have a C-suite member with a formal mandate to drive analytics across technology, people, process, data and strategy. Organizations must evolve beyond unstructured and siloed analytical capabilities and rather develop and maintain a thriving ecosystem for analytical capabilities to interact with rest of the business units in order to produce tangible results.

One of the approaches is to design a hub-and-spoke model wherein the analytics team is established as a Centre of Excellence (COE) - interfacing with various business units to solve specific use cases. One of the advantages of having such a model is that, business units can come together to solve problems which will be inter-functional and will also require them to share their departmental data. For instance, in order to have a 360-degree understanding of a product's performance, the COE may have to obtain data from the marketing department to identify the cost invested on various marketing channels, obtain data from the customer support department in terms of the number of calls they receive from the Customers for various products, obtain data from organization's social media page etc.

Fig. 3 provides an organizational

structure for a typical analytics COE. The analytics champion(s) should report directly to the CEO, CMO, COO or CFO. This serves dual purpose. From the business perspective, the C-suite executives will be able to see the transformational change resulting from the use of analytics. From an analytics team's perspective, it helps them to keep the business impact in mind without getting lost in the technical intricacies. Additionally, this approach also has the advantage of earning the confidence of top management at an initial stage, compared to fighting the way upward, by overcoming inter-departmental conflicts. It can even smoothen the cultural transition towards an insight-driven organization.

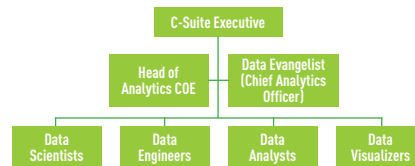


Fig. 3: Typical Organisational Structure of an Analytics COE

Further, every analytics project should be communicated to the wider organization during review meetings or through periodic newsletters so that, every department is aware about how the organization is leveraging the analytics services being provided by the COE to solve some of their challenges or to gain insights from data which were not directly visible to them. For instance, the marketing team may benefit immensely if they are given insights on how a campaign is performing across various age groups in different locations. This feedback can possibly help them to improve their target marketing strategies.

One of the effective ways to spread analytics to every department is by having one of their representatives working intermittently in the COE as a domain expert. This domain expert

can be a business unit head. This will help COE to communicate what it does to the wider department through this departmental nominee rather than the former explaining their work! This approach can even facilitate in improving the explainability of analytics as a technology.

Finally, every successful project which used analytics at any stage of the decision making should be celebrated. The concerned stakeholders and the COE should even communicate the scale of impact; for instance in terms of increasing revenue, improving the sales, increasing the customer retention, or reduction in the time taken to make a decision etc.

Conclusion

Monetizing data present within an organization will require C-Suite executives to embrace the fact that data is an asset rather than a liability. Also, it is equally important to align analytics deliverables to tangible business goals. Finally, solving business issues by collaborating with a centralized analytics COE can not only overcome inter-departmental conflicts but also, expedite the organization's journey towards being an insight-driven organization.

References

- [1] 5 Trends Emerge in the Gartner Hype Cycle for Emerging Technologies, 2018, <https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018/>
- [2] Insight Driven Organization Survey: Benchmarking your Analytics Journey, Deloitte, 2018 <https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/technology/deloitte-uk-tech-ido-survey-report-2018.pdf>

Disclaimer :

All views expressed by the author in this article are strictly personal.

About the Author



Binesh Nair is a Senior Consultant in Analytics and Information Management in Deloitte, Ireland. He has over 6 years of experience in Consulting and Analytics; collaborating with big and small businesses across diverse domains. He was a Government of Ireland Scholar in 2014/15 and has published 20+ papers in applied data analytics.



A Framework for Software Fault Prediction using Decision Tree Based Machine Learning

► Preeti Gulia

Department of Computer Science and Applications, MDU, Rohtak, India
Email: preeti.gulia81@gmail.com

► Palak

Government College, Sector 9, Gurugram, India
Email: palak.aug6@gmail.com

Software plays an important role in our daily life. The use of smart real time devices has increased dramatically over last decade which in turn raising the need for fault proof, highly reliable software. The quality expectations of end user are very high in this era of technology that demands software to be robust and failure proof. Manual methods of predicting and identifying faults in complex systems may not guarantee fault free system due to some inherent limitations and they often are time consuming. The main expectation from reliable software is the minimization of the number of failures that occur when the program runs. Software defect prediction is important for identification of defect-prone parts of software. One approach in achieving this is to dynamically assess the modules in the synthesized code using software defect prediction techniques based on machine learning. In this paper, decision tree based machine learning technique is applied to assess its efficiency on three different performance parameters.

Keywords: Machine Learning; Software; Fault Prediction; decision tree.

I. Introduction

A well planned software engineering process aims in delivering fault free software that meets the user specifications and are easy to maintain. Faults in the software are the conditions where system does not perform as expected. Over years software engineering has improved a long way to deliver high quality software but still there are chances of error at the various stages of SDLC (software development life cycle). Software testing is an important and unavoidable phase to detect and remove defects. It should start early in stages of SDLC so that errors do not propagate from one stage to another. Inadequate software testing may result in system failure at later stages of implementation. A single defect can cause the entire system failure and most of the time they drop the quality of the software system drastically. Important activities of bug prevention are software testing, inspection, formal verification and software defect prediction. The growing complexity of software resulted in the need to describe it at different levels

of abstraction, and, in addition to this, the notion of software architecture has developed. Software fault prediction is a discipline that predicts whether software will be fault prone in future by using historical background of software product lines. In software engineering, early detection of defective portions of the software can help the software developers and engineers in proper allocation of limited resources in testing and maintenance phases of the software development.

Machine learning (ML) includes training of computer systems to recognize patterns and behavior based on historical data to make decisions and to solve computational problems without using any explicit aid. ML is considered as an inseparable aspect of artificial intelligence. ML uses training data to teach machine how to respond to the real life problems and imitate the behavior of human or the living organisms. Machine learning is about predicting the future based on the past. Implementation of machine learning is broadly classified into two

types: supervised and unsupervised learning [1]. In supervised learning, training data contains both input and output. So the results can be easily monitored and compared to the desired output. The whole system can be supervised and modifications are done to minimize the gap between actual and desired output. On the other hand, unsupervised learning is used to train such systems when output is not known in advance. Errors in this case are mitigated by some random adjustments in the systems. Anomaly detection and clustering are some of the examples where unsupervised learning is used. In between supervised and unsupervised learning, there is one more learning paradigm called semi-supervised learning which takes the advantage of both. It may sometimes also called as reinforcement learning in which despite of knowledge of desired output, the error signal is based on some random value.

II. Related Work

Machine learning has attracted many researchers over last few

decades due to its myriad applications in various fields of engineering and technology. Software fault prediction is one such area of application of ML tools and techniques. Various ML techniques like neural networks, fuzzy systems, decision tree, naïve Bayes etc. are seemingly utilized for prediction of future fault prone software. A vast literature is available online and offline in this regard. A few important researches are reviewed in this section.

P. Bal et. al. in [2] used extreme learning machine (ELM) for cross project defect prediction. Cross projects are those projects which are based on historical data of other projects. Their focus is on completeness of fault prediction for different datasets. One more important research in this league is given in [3] where authors applied unsupervised classifier for cross project defect prediction. They also compared the results with supervised classifiers.

Authors in [4] used various machine learning techniques of bug prediction namely artificial neural network, Naïve Bayes and Decision tree. They tested their results on three real datasets and concluded that decision tree performs best out of the three. Feature selection is an important aspect in applying machine learning.

A. Jakhar et. al. in [5] proposed a feature selection model for defect prediction on NASA software engineering data sets. They utilized seven feature selection techniques and eight classifiers for this model. Authors in [6] applied machine learning techniques on python based software to assess whether the models used for C/C++ and java are effective for python or not. They found the approach appropriate for python based programs with lower false positive rate. Android software is also gaining popularity in past few years. Researchers find it attractive to apply ML techniques of fault prediction in Android based systems.

One such effort is given in [7] where author used object-oriented metrics for predicting defective classes using 18 machine- learning techniques. The proposed framework has been applied to seven application packages of well known, widely used Android operating system viz. Contact, MMS,

Bluetooth, Email, Calendar, Gallery2 and Telephony.

Authors in [8] reviewed three machine learning based techniques for software fault prediction namely Support Vector Machine (SVM), Radial Basis Function Network (RBF) and Adaptive Boost Network. A new model for bug prediction can be seen in [9] where authors predicted the bug proneness index using marginal R square values after applying multiple regression.

Authors in [10] summarized some important open issue in applying fault prediction techniques. Some of them are finding the relationships between attributes and faults, standard specification, cross project issues, lack of general framework etc. Neural network and fuzzy systems are two ML techniques that always attracted researchers of various fields. A comparison is carried out between such soft computing based techniques to assess their relevance in fault prediction [11]. The problem with neural network is that they require deep understanding with the system so it becomes difficult to represent the problem to the network. One cost effective neural network based model is presented in [12].

Authors in [13] utilized Bayesian Regularization (BR) for fault prediction. Their results signify that the software fault prediction model using BR technique provide better accuracy than Levenberg-Marquardt (LM) algorithm and Back Propagation (BPA) algorithm. Authors in [14] utilized various data mining approaches for fault and effort prediction. Similar effort is done in [15] for software static defect prediction. According to their experimental results Naïve Bayes performs better than other data mining techniques.

III. Decision Tree

A decision tree can be used to visually represent decisions and decision making. Each path from decision node to leaf node gives us a rule. It is a type of supervised learning in which class labels are already known. It can be used for both classification and regression. A decision tree can be built using greedy approach. For attaining high accuracy the height of the tree should be small. Decision Tree models are created using

2 steps: Induction and Pruning. Decision Tree algorithms are referred to as CART or Classification and Regression Trees. We can represent any boolean function on discrete attributes using the decision tree. Fig. 1 shows decision tree for OR gate.

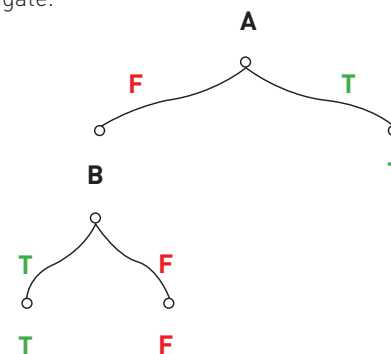


Fig. 1: Decision Tree for OR gate

IV. Results and Analysis

To assess the performance to decision tree based machine learning approach, three well known performance measures are taken into consideration i.e. accuracy, precision and recall that are derived from confusion matrix. Confusion matrix is a table having counts for true positive, true negative, false positive and false negative. The aim is to achieve high true positive and true negative. Accuracy is the proportion of true results among the total number of examined instances. Precision is calculated as the count of true positive predictions divided by the total number of positive predictions. Recall is calculated as the count of true positive predictions divided by the total number of positives. Table 1 shows the result of applying Decision Tree to a sample dataset with 109 records containing number of faults and number of test workers on daily basis. The results show that decision tree is a good choice in the field of software fault prediction.

Table 1: Performance Analysis of Decision Tree

Performance Measure	Decision Tree Performance
Accuracy	0.970
Precision	0.990
Recall	0.999

V. Conclusion

Software industry is growing at a very fast pace with the growing need of smart devices. Faulty software never satisfies end user expectation and degrades organization goodwill. Inadequate testing may result into software failure at later stages. Software fault prediction techniques have proved a boon to the developers where historical data is used to predict future faults in related software. This paper analyses the performance of decision tree based software fault prediction approach which is a well known machine learning technique. Three performance parameters are taken into consideration to assess the effectiveness of decision tree. The results show that decision tree is a good choice in the field of software fault prediction.

References

- [1] T. Oladipupo, "Types of Machine Learning Algorithms," *New Adv. Mach. Learn.*, 2012.
- [2] P. R. Bal and S. Kumar, "Cross Project Software Defect Prediction using Extreme Learning Machine : An Ensemble based Study," in *Proceedings of the 13th International Conference on Software Technologies (ICSOFT 2018)*, 2018, pp. 320-327.
- [3] F. Zhang, Q. Zheng, Y. Zou, and A. E. Hassan, "Cross-project Defect Prediction Using a Connectivity-based Unsupervised Classifier," 2016.
- [4] A. Hammouri, M. Hammad, M. Alnabhan, and F. Alsarayrah, "Software Bug Prediction using Machine Learning Approach," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 2, pp. 78-83, 2018.
- [5] A. K. Jakhar and K. Rajnish, "Software Fault Prediction with Data Mining Techniques by Using Feature Selection Based Models," *Int. J. Electr. Eng. Informatics*, vol. 10, no. 3, pp. 447-465, 2018.
- [6] B. Wójcicki and R. Dąbrowski, "Applying Machine Learning to Software Fault Prediction," *e-Informatica Softw. Eng. J.*, vol. 12, no. 1, pp. 199-216, 2018.
- [7] R. Malhotra, "An empirical framework for defect prediction using machine learning Techniques with Andriod Software," *Appl. Soft Comput. j*, 2016.
- [8] K. Manchanda, H. Wadhwa, and H. Kaur, "A Systematic Review Based On Machine Learning Techniques for Software Defect Predication," *Int. J. Eng. Dev. Res.*, vol. 4, no. 2, pp. 1290-1292, 2016.
- [9] S. Puranik, P. Deshpande, and K. Chandrasekaran, "A Novel Machine Learning Approach for Bug Prediction," *Procedia - Procedia Comput. Sci.*, vol. 93, no. September, pp. 924-930, 2016.
- [10] I. Arora, V. Tatarwal, and A. Saha, "Open Issues in Software Defect Prediction," *Procedia - Procedia Comput. Sci.*, vol. 46, no. Ictict 2014, pp. 906-912, 2015.
- [11] E. Erturk and E. Akcapinar, "A comparison of some soft computing methods for software fault prediction," *Expert Syst. Appl.*, vol. 42, no. 4, pp. 1872-1879, 2015.
- [12] Ö. Faruk, "Software defect prediction using cost-sensitive neural network," *Appl. Soft Comput.*, vol. 33, pp. 263-277, 2015.
- [13] S. Kumar and R. Kumar, "Design of Software Fault Prediction Model Using BR Technique," *Procedia - Procedia Comput. Sci.*, vol. 46, no. Ictict 2014, pp. 849-858, 2015.
- [14] M. Julie, J. D. F. Enric, D. Karel, B. Bart, M. David, J. Moeyersoms, and E. Junqu, "Comprehensible software fault and effort prediction : a data mining approach," *J. Syst. Softw.*, vol. 100, pp. 80-90, 2015.
- [15] A. H. Yousef, "Extracting software static defect models using data mining," *Ain Shams Eng. J.*, vol. 6, no. 1, pp. 133-144, 2015.

About the Authors



Dr. Preeti Gulia is currently working as Assistant Professor at Department of Computer Science & Applications, M.D.University, Rohtak, India. She is serving the Department since 2009. She earned her doctoral degree in 2013. She has published more than 65 research papers and articles in journal and conferences of National/ International repute including ACM, Scopus. Her area of research includes Data Mining, Big Data, Machine Learning, Deep Learning, IoT, Software Engineering. She is an active professional member of IAENG, CSI and ACM. She is also serving as Editorial Board Member Active Reviewer of International/ National Journals. She has guided one research scholar as well as guiding four Ph.D. research scholars from various research areas.



Ms. Palak is presently serving as Assistant Professor at Government College, Sector 9, Gurugram. She has earned M.Tech. degree from Department of Computer Science & Applications, M.D.University, Rohtak, India in 2014. Currently she is pursuing Ph.D. under the supervision of Dr. Preeti Gulia, Assistant Professor, M.D.University, Rohtak, India. Ms. Palak is a receiver of Junior Research Fellowship from UGC for the period of 2017-2018. Her area of research includes Soft Computing, Metaheuristics, Machine Learning, Software Testing. She has authored many research papers of International indexing including Scopus.

Managing Credit Risk in Banking Application using Data Analytics – A Bird View

► Archana Sasi

Assistant Professor, Department of CSE, School of Engineering, Presidency University, Karnataka. Email: archanasasi@presidencyuniversity.in

Data Analytics - Introduction

Data analytics play a decisive role in making decisions that motivate organizations to move forward, helping to improve efficiency, increase profits and thereby achieve their business goals using massive chunks of data. Much effort should be required at various levels to make decisions so the core idea behind the process of data analytics is to discover, represent and convey important data insights to achieve better decision-making.

Data Analytics Applications – Real time

Data analytics applications are growing considerably as a result of innovations from different fields. It have been used in various areas such as transport, fraud and risk detection, risk management, banking and finance, health care and customer interactions.... In general, risk management is classified into five segments, namely market risk, capital adequacy, liquidity risk, operational risk, and credit risk. In this article, we focus on credit risk management in banking applications. The Schematic diagram of risk management is represented using Fig. 1.



Fig 1. Structure of Risk Management

Credit Risk Management Framework in Practice

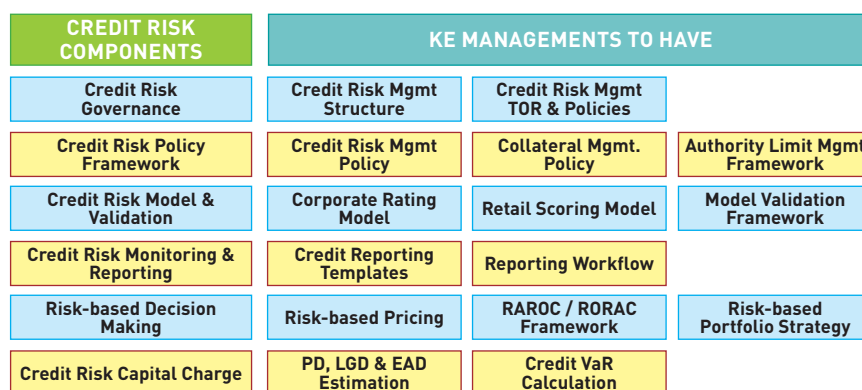


Fig. 2 : Key Risks in Banking Application

Credit Event and Credit Risk – A Definition

Any loss that a lender can expect from the loan and borrowing process is called a credit event. Credit risk is primarily related to the loss that can typically occur due to some event related to credit. It refers to the likelihood of loss that the borrower may result from failing to make the necessary payments in accordance with agreed terms.

Credit Risk Management Framework – In Banking Application

Credit risk management plays an important role in banks to improve their banking and performance operations. It also protects the bank's resources by prioritizing risks and by planning to deal with each possibility. The overall architecture for key risks in banking application is shown using Fig. 2.

Role of Data Analytics in Credit Risk Management

Data analytics plays a crucial role in managing credit risk in the banking sector. Analysis of credit risk provides lenders with a complete customer profile and also helps them to expect customer behavior. Using these

analytical techniques, lenders can save money, time, and resources in order to target the right customers and expect the risks involved. In addition, the risk level for those customers with no credit history can be observed. Credit risk analyst is the one who does this analysis work for banks or other institutions. Their main task is to evaluate applications for loans and determine who will pay the loans back. Diagram 3 shows how credit risk is managing in banks.



Fig. 3: Credit Risk is management in Banking Sector

Conclusion

Credit risk management traditionally checks the debt-to-equity ratio and analyzes where the loan is concentrated across portfolios, geography and other attributes so that the bank can decide whether to take more risk or take less risk. There are some best practices that can be implemented faster and more efficiently with the emergence of data analytics. The need of this technology is used by banks use credit risk management with accurate and acceptable parameters

to retain exposure due to credit risks. It is directly applied on real time banking operation such as integration of systems, data quality and modeling. The outcome of this data analytics in credit risk management are applied over banking sector and it has been recognized at a higher level as a result, most banks have started to earn profits.

Reference

- [1] <https://www.google.com/search?q=credit+risk+management+in+banks&hl=en-IN&authuser=0&source=lnms&tbn=isch&sa=X&sqi=2&ved=0ahUKEwiVOP>

vhh-PhAhVmKLkGHQvnAXUQ_AUIECgD&biw=1366&bih=608#imgdii=gtPJSEPeULBZ9M:&imgsrc=mSYAPkY2YdK__M:

- [2] https://www.google.com/search?q=credit+risk+management+framework+for+banks&source=lnms&tbn=isch&sa=X&ved=0ahUKEwialYbMvOPhAhXH7HMBHeQNB28Q_AUIDigB&biw=1366&bih=608#imgsrc=gPhDHpuznYjE3M
- [3] https://www.google.com/search?q=structure+of+risk+management&source=lnms&tbn=isch&sa=X&ved=0ahUKEwijtpTmvOPhAhXEX3wKHWE3APIQ_AUIDigB&biw=1366&bih=608#imgsrc=M0MdlycEoi_rM

About the Author



Archana Sasi received her B.Tech and M.Tech in Computer Science and Engineering and Computer and Information Science from College of Engineering Chertala in 2012 and 2014 respectively. She is currently working as Assistant Professor in Computer Science and Engineering Department at Presidency University Bangalore, Karnataka. Her research area includes Data Mining, Machine Learning and Big Data Analytics.

Congratulations



Prof. Sinha appointed as International Development Consultant

With a vision of holistic development in line with UN 2030 goals of sustainable peace and development, Republic of South Sudan has appointed Prof. (Dr.) Ripu Ranjan Sinha Vice-Chairman cum Elected Chairmen, Computer Society of India, Jaipur Chapter as consultant for socio- economic development Republic of South Sudan. H.E Pro .Gen. James.Wani Igga, Vice President , Republic of South Sudan extended his best wishes to the president and members of Council for the successful implementation of Holistic development mission 2030. The Republic of South Sudan has empowered Prof. Dr. Ripu Ranjan Sinha to frame the policy for Socio -economic development and also promised to provide all mandate necessities for attaining the goals of sustainable peace and development .Provided that in present scenario Republic of South Sudan is one of the developed nation among East African Countries and also one of the prominent nation in the global community for investment by 2030 because of emerging market trends and clamor for ICT , science and technology.Prof. Dr. Ripu Ranjan Sinha expressed his gratitude for this distinctive honor and unique part of this great movement

Handwritten Digits Recognition using Convolutional Neural Network

► **Amit Choudhary**

Associate Professor, Department of Computer Science,
Maharaja Surajmal Institute, New Delhi
Email: amit.choudhary69@gmail.com

► **Savita Ahlawat**

Reader, Department of Computer Science and Engineering,
Maharaja Surajmal Institute of Technology, New Delhi
Email: savita.ahlawat@gmail.com

1. Introduction

Convolutional Neural Network (CNN) is basically a variant of Multi-Layer Perceptron (MLP) network and was used for the first time in 1980. The computing in CNN is inspired by the activity of human brain. Humans perceive or identify the objects using visuals. We (humans) generally train our child to recognize various objects by showing him many pictures of that object. This makes him to identify or make a prediction for the objects he has never seen before. The CNN works in the same fashion and is popular for analyzing visual imagery. Some of the well-known CNN architectures are GoogLeNet (22 layers), AlexNet (8 layers), VGG (16-19 layers), and ResNet (152 layers). The CNN integrates the feature extraction and classification steps and requires minimal pre-processing and feature extraction efforts. The CNN can extract affluent and interrelated features automatically from images. Moreover, CNN can provide considerable recognition accuracy in case of availability of small training data. It eliminates the necessity of collecting prior knowledge and design details of features. The major advantage of the CNN model is that it exploits the topological information present in the input and is invariant to basic transformation like rotation, translation etc. CNN can also scale well for high resolution images.

The convolutional neural network is a specific type of deep neural network having wide application in image classification, object recognition, recommendation system, signal processing, natural language processing, computer vision and face recognition. Some CNN variants are - Autoencoders, Recurrent Neural Networks (RNN), Recursive Neural

Network, Deep Belief Networks and Deep Boltzmann Machines etc. The ability to automatically detect the important features of an object (here object can be an image, a handwritten character or a face etc.) without any human supervision or intervention makes them (CNN) efficient in comparison to its predecessors (MLP etc.). In other words, the high capability of hierarchical feature learning results in a highly efficient CNN.

The main applications are vehicle number plate recognition, road side object detection, postal letter sorting services, Scene Labelling, CTS scanning and historical document preservation in the archaeology department, Human Pose Estimation, Computer Vision, old documents automation in libraries and bank etc. All these areas deal with large image databases and hence demands high recognition accuracy, lesser computational complexity and consistent performance of the recognition system.

2. Convolutional Neural Network Architecture Overview

The CNN computation is inspired by the visual cortex in humans. The visual cortex is a part of the brain that processes the information forwarded from the retina. It processes the visual information and is subtle to small sub-regions of the input. A basic convolutional neural network comprises of three components namely the convolutional layer, pooling layer and the output layer. Neurons in one layer connect with some of the neurons present in the next layer making the scaling easier for the higher resolution images. The operation of pooling or sub-sampling can be used to reduce the dimensions of the input. In the CNN model, the input image is considered as

a collection of small subregions called the "receptive fields". A mathematical operation of the convolution is applied on the input layer which emulates the response to the next layer. The response is basically a visual stimulus.

The input data is loaded and stored in the input layer. This layer specifies the height, width, and number of channels (RGB information) of the input image. The hidden layers are the backbone of CNN architecture. They perform a feature extraction process where a series of convolution, pooling and activation functions are used. The distinguishable features of visual images are detected at this stage. The convolutional layer is the first layer placed above the input image. It is used for extracting the features of an image. It introduces non-linearity through neural activation function. A receptive field is calculated in CNN, which is a small region of an input image that can affect a specific region of the network as shown in Fig. 1.

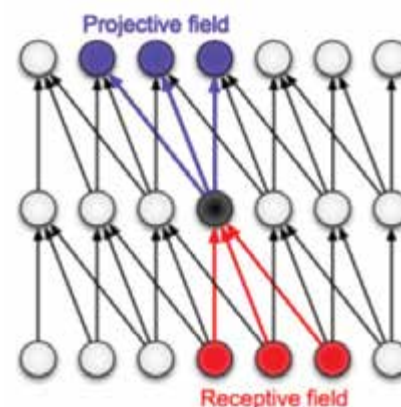


Fig.1: Receptive Field and Projective Field

It is also one of the important design parameters of the CNN architecture

and helps in setting other CNN parameters. It has the same size as the kernel and works in similar fashion as the human eye works for producing the sharp central vision. The receptive field is influenced by striding, pooling, kernel size, and depth of CNN. Stride is another parameter used in the CNN architecture. It is defined as the step size that the filter moves every time. The stride value of 1 indicates filter sliding movement pixel by pixel. A larger stride size shows less overlapping between the cells. The working of kernel and stride in convolution layer has been presented in Fig. 2.

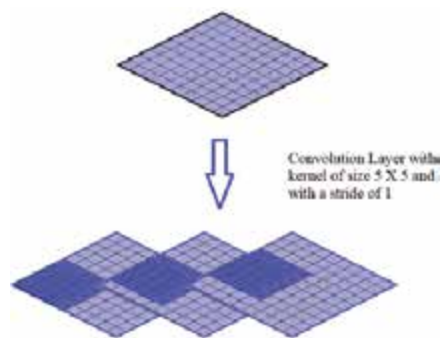


Fig. 2 : The Representation of Kernel and Stride in Convolutional Layer

3. Experimental Setup

To accomplish the task of handwritten digit recognition, a model of the convolutional neural network is developed and analyzed for different parameters like recognition accuracy and processing time. The proposed work investigates two variants of the CNN architecture, three layers architecture (CNN_3L) and four layers architecture (CNN_4L). A total of six cases (case 1 to case 6) have been considered for the CNN with three-layer architecture and five cases (case 1 to case 5) for the CNN architecture with four layers. All the cases differ in number of feature maps, stride sizes, padding, dilation, and received receptive fields.

All the experiments have been performed in the MATLAB 2018b on a personal computer with Windows 10, Intel (R) Core (TM) i7-6500 CPU (2.50 GHz), 16.00 GB memory and NVIDIA GPU. The MNIST database has been involved in training and testing models. The standard MNIST handwritten digit database has 60000 and 10000 normalized digit images in its training

and testing dataset respectively. Some of the sample images from MNIST database has been shown in Fig. 3.

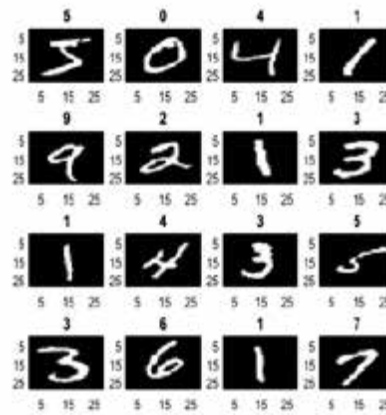


Fig. 3 : Sample MNIST Handwritten Digit Images

Data pre-processing plays an important role in any recognition process. For shaping the input images in a suitable form for segmentation, the pre-processing methods such as scaling, noise reduction, centering, slanting, skew estimation and many more have been used. In general, many algorithms show improvement in performance after the data has been normalized and whitened. One need to work with different algorithms in order to find out the exact parameters for data pre-processing. In the present work, the MNIST images have already been size normalized into a fixed image of size 28 x 28.

The convolution layer with 28 kernel/filter/patch and kernel sizes of 5x5 and 3x3 are used here to extract the features. Every patch contains the structural information of an image. In Fig. 2, a convolution operation is computed by sliding the filter of size 5x5 over the input image. This layer is responsible for transforming the input data by using a patch/filter of locally connecting neurons from the previous layer.

4. Results and Discussion

The receptive field (RF) is calculated based on kernel size (k), dilation (d), padding (p), stride (s), input size (i/p) and output size (o/p) of the feature map. The value of the receptive field, when close to the input size, observed good recognition accuracy for all the feature maps. On the other hand, a large gap

between receptive field and input size observed poor recognition accuracy. Training with the "Adadelta" optimizer can further increase the accuracy of the classifier.

The experimental results of the MNIST handwritten digit dataset using different parameters of CNN_3L and CNN_4L architectures are recorded and analyzed. A sample image from MNSIT dataset is represented as a 1-D array of 784 (28 x 28) float values between 0 and 1 (0 stands for black, 1 for white). The findings confirm the role of different architecture parameters on the performance of a recognition system. The training parameter used here has a learning rate of 0.01, the maximum epoch counts of 4. The highest recognition accuracy achieved in Case 3 with CNN_3L architecture having three layers is 99.76% for the feature map 12-24-32. In Case 5, the CNN_4L architecture with four layers achieved the highest recognition accuracy of 99.76 % for the feature map 12-24-28-32. The stochastic gradient descent with momentum [SGDm] can also be used to achieve higher performance.

5. Conclusion

The CNN architecture avoids complex pre-processing, costly feature extraction procedure, complex classifier combination approach and proved to be better than the traditional pattern recognition approaches. The values of different architectural parameters should be selected wisely. The increased number of layers might cause overfitting and consequently can influence the recognition accuracy. The problem of the overfitting can be avoided by finding out optimal values using trial and error method or under some guidance. The concept of dropout may be used to solve the problem of overfitting in which we can stop some randomly selected neurons (both hidden and visible) from participating in the training process. Basically, dropout is the weight regularization technique and most preferred in the larger networks to achieve better outcomes. Generally, a small value of dropout is preferred otherwise network may be under learning.

References

- [1] Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A Deep Convolutional Encoder-Decoder

- Architecture for Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12), 2481-2495.
- [2] Fayek, H. M., Lech, M., & Cavedon, L. (2017). Evaluating deep learning architectures for speech emotion recognition. Neural Networks, 60-68.
 - [3] Guimin, L., Qingxiang, W., Lida, Q., Xixian, H. [2018]. Image super-resolution using a dilated convolutional neural network, Neurocomputing, Volume 275, 2018, Pages 1219-1230, ISSN 0925-2312, <https://doi.org/10.1016/j.neucom.2017.09.062>
 - [4] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84-90.
 - [5] Liang, T., Xu, X., & Xiao, P. (2017). A new image classification method based on modified condensed nearest neighbor and convolutional neural networks. Pattern Recognition Letters, 94, 105-111.
 - [6] Qu, X., Wang, W., Lu, K., & Zhou, J. (2018). Data augmentation and directional feature maps extraction for in-air handwritten Chinese character recognition based on convolutional neural network. Pattern Recognition Letters, 111, 9-15.
 - [7] Sainath T.N., Kingsbury B., Saon G., Soltau H., Mohamed A.R., Dahi G., et al. [2015]. Deep convolutional neural networks for large-scale speech tasks Neural Networks, 64, 39-48.
 - [8] Scherer, D., Muller, A., & Behnke, S. (2010). Evaluation of pooling operations in convolutional architectures for object recognition. In Proceedings of 20th International Conference on Artificial Neural Networks (ICANN), 92-101.
 - [9] Shi, Z., Ye, Y., & Wu, Y. (2016). Rank-based pooling for deep convolutional neural networks. Neural Networks, 83, 21-31.
 - [10] Zhang, Z., Wang, H., Liu, S., & Xiao, B. (2018). Deep Contextual Stroke Pooling for Scene Character Recognition. IEEE Access, 6, 16454-16463.

About the Authors



Dr. Amit Choudhary is currently working as an Associate Professor in the Department of Computer Science at Maharaja Surajmal Institute, New Delhi. He has done MCA (Kurukshetra University), M.Tech. (Delhi Technological University), & Ph.D. (Maharshi Dayanand University) in Computer Science and holds 16 years of experience. He has published around 35 papers in international journals and conferences. He is a reviewer of many National & International Journals and Conferences. His research interests are Pattern Recognition, Soft Computing, Machine Learning, Big Data Analytics & Computer Vision. He is a member of CSI.



Ms. Savita Ahlawat is currently working as a Reader in the Department of Computer Science and Engineering at Maharaja Surajmal Institute of Technology, New Delhi. She has done B.Tech (Maharshi Dayanand University), M.Tech. (GGSIIP University) & pursuing Ph.D. (Maharshi Dayanand University) in Computer Science and Engineering. She holds 15 years of experience. She has published around 25 papers in international journals and conferences. Her research interests include Pattern Recognition, Machine Learning, Data Science & Computer Vision. She is a member of ISTE.

About the Guest Editor



Prof. Nasib Singh Gill is the senior most Professor of Department of Computer Science & Applications, M. D. University, Rohtak and is working in the Department since its inception in the year 1990. He is also currently working as the founder Director of Digital Learning Initiatives & Monitoring Cell, M. D. University, Rohtak. Professor Gill earned his Doctorate in Computer Science in the year 1996 under the supervision of Professor P.S. Grover of Delhi University. Prof. Gill carried out his Post-Doctoral research at Brunel University, West London during 2001-2002. Dr. Gill is recipient of Commonwealth Fellowship Award of British Government for the Year 2001. Besides, he also has earned his MBA degree.

He has also held several major responsibilities in past too such as Head of Department; Director, University Computer Centre; Director, Directorate of Distance Education; Director, MDU Alumni; Director, Public Relations; University Nodal Officer, All India Survey on Higher Education (AISHE), MHRD, Govt of India at M.D. University Rohtak and Software Engineer in Haryana State Electronics Development Corporation Limited (HARTRON), Chandigarh prior 1990

He has widely travelled abroad for several academic and research assignments including countries like UK, USA, France, etc. Professor Gill has guided so far 9 Ph.D. scholars as well as guiding about 7 more scholars presently. He has published more than 260 research papers in National & International Journals, Conference Proceedings, Bulletins, Edited Books, and Newspapers. He has been awarded with 'Best Paper Award' by Computer Society of India in the year 1994 for contributing the best paper "A New Program Complexity Measure" in their Journal. He has authored seven books including some popular books on 'Software Engineering', 'Handbook of Computer Fundamentals' and 'Computing Fundamentals and Programming in C'.

Professor Gill is a Senior Member of IACSIT as well as a fellow of several professional bodies including IETE (The Institution of Electronics and Telecommunication Engineers), Computer Society of India (CSI). He has been serving as Editorial Board Member, Guest Editor, Reviewer of International/National Journals and a Member of Technical Committee of several International/National Conferences.

Role and Impact of Big Data Analytics in Health Care Systems

► Mamoon Rashid

Assistant Professor, School of CS & Engg.,
Lovely Professional University, India.
Email: mamoon873@gmail.com

► Kulvinder Singh

B.Tech Student, School of Computer Science &
Engineering, Lovely Professional University, India.
Email: kulvindersingh97@gmail.com

► Harjeet Singh

Associate Professor, Dept. of Computer Science,
Mata Gujri College, Fategarh Sahib, India.
Email: zrjeet@gmail.com

► Neeraj Mogla

Visual Insights & Analytics Manager,
Nike Inc. Global Head Quarters, USA.
Email: moglaneeraj@gmail.com

► Vishal Goyal

Professor, Department of Computer Science,
Punjabi University, Patiala.
Email: vishal.pup@gmail.com

1. Introduction:

Presently volumes of data is generated in Health Care Systems in the form of test reports, clinical prescriptions, billing reports and laboratory tests [1]. It is a real challenge in Health Care Systems to process this huge amount of data in different formats. Moreover it is important to predict and analyze this Healthcare data with faster rates without being stored in traditional ways. Big Data Analytics is one such approach which addresses these issues in Health Care systems. With the help of Big Data pipelines, it is quite possible now to reduce the costs for various treatments, detect and predict diseases at earlier stages and minimize the human error in terms of diagnosing among patients.

2. Definition of Big Data for Health Care Systems:

Big Data refers to the voluminous information which is processed and analyzed by various technologies with less computation time [2]. Health Care Big Data is the data in huge quantities which is available in health databases like Electronic Health Records (EHR). The overview of Big Data in Health care is summed up in Fig. 1.

In Healthcare systems, it will make use of patient data and provide the ways to detect such individuals from different kind of diseases on low cost basis [3]. In today's Health Care Systems, medical experts expect best possible understanding related to patient history that too as early as possible. This demands huge amounts of medical data related to patients whose processing needs to be done with less costs and time. Health Care

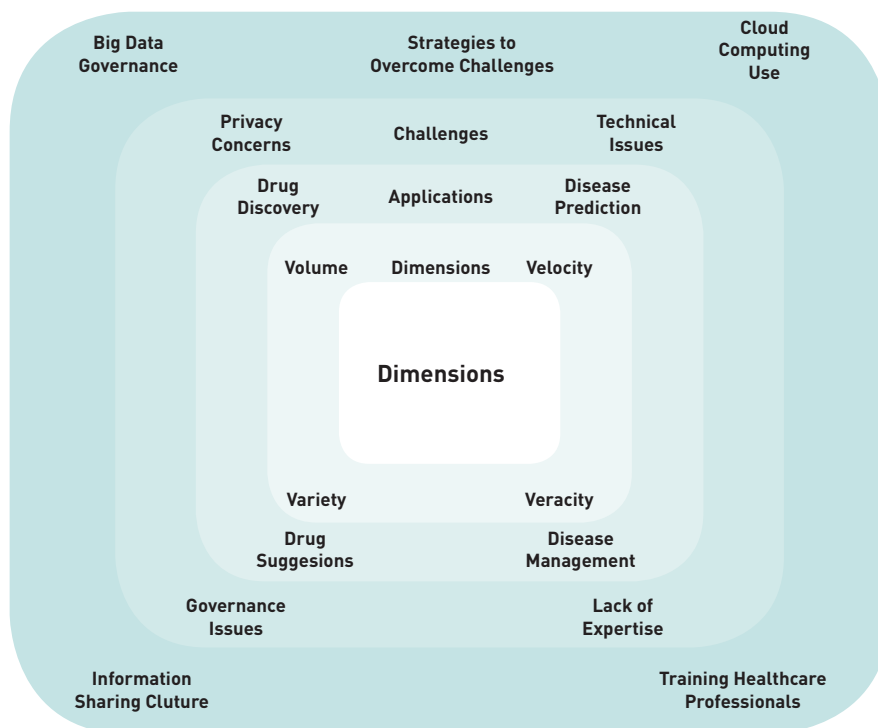


Fig. 1 : Overview of Big Data Healthcare Data

Analytics in terms of Big Data pipelines is the possible solution of these issues where all patient data is collected and stored effectively and processed it to get critical insights that help medical doctors to take decisions for better medical care.

3. Architecture of Big Data Analytics for Health Care Systems:

The architecture of Big Data Analytics for Health care Systems consist of five layers: Data Layer, Data Aggregation layer, Analytics Layer,

Information Exploration Layer and Data Governance Layer [4]. The layers in Healthcare Analytics architecture are shown in Figure 2.

Data Layer: This layer is responsible for collecting data from different sources of medical laboratory equipment's and Electronic Health Records. This collected data is of three formats: Structured, semi- structured and unstructured data. Structured data is having fixed format and is easily stored, processed and accessed.

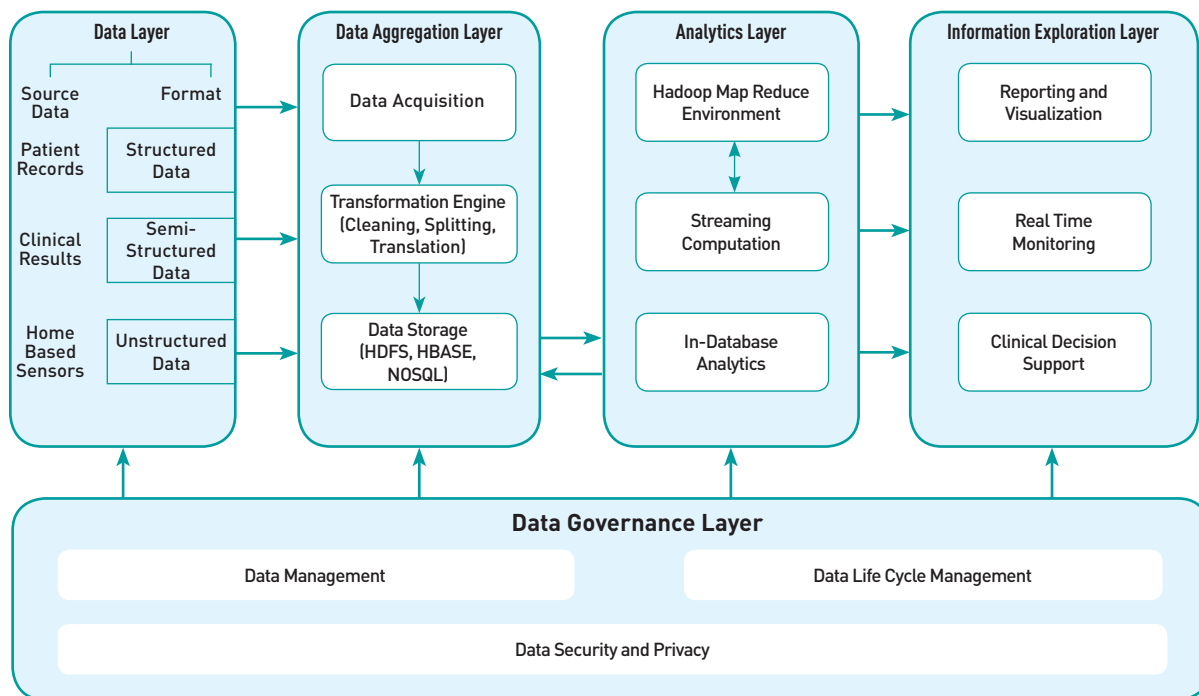


Fig. 2 : Architecture of Big Data Analytics for Health Care Systems

Structured data is always following particular order as in row and column format and always results into ordered output. All traditional databases containing data in row column format belong to this category. **Unstructured** data is usually huge data which is not in organized manner. This kind of data remains usually unknown and poses numerous challenges while processing for valuable insights as output. Data in the form of images, audio, video and sensor based data belong to this category. **Semi-structured** data usually contains both the forms but remain undefined. Representation of data in terms of XML files belong to this category [5].

Data Aggregation Layer: This layer is responsible for the storage and preprocessing operations of data. The operations performed on this layer are the extraction of data, its transformation in particular formats and storage on file systems. Apache Hadoop is one such Big Data platform which is used for storage of big data. Hadoop Distributed File System (HDFS) is the component which is used for the storage of big data in various formats.

Analytics Layer: This layer is

responsible for the processing of stored data on big data environment. Apache Hadoop provides programming model of Map Reduce which is required for efficiently computing this medical big data. Data is divided into multiple nodes on storage cluster of HDFS in terms of blocks and then maps are sent on each node for the processing of data on nodes. The intermediate results achieved on nodes are aggregated with the help of reducer which gives the final result. Both Map and Reducer are programming scripts which are running on nodes itself to process the big data with least computational times.

Information Exploration Layer: This layer is responsible for visualizing results and monitoring. This layer will provide insights about the complex and huge data inputted to Big Data pipeline on which basis diagnosis and decisions are to be taken. This layer provides the summarization of results for the input data.

Data Governance Layer: This layer provides support to all layers and is responsible for data management and its privacy. This layer is integrated with all remaining layers of architecture and governs the Big Data pipeline for all

kinds of operations.

4. Need of Big Data Analytics in Health Care Systems:

The adoption of Big Data Analytics in Health Care Systems is due to the following requirements:

Storage of Healthcare Data: Due to the digitization of healthcare data, the amount of Healthcare data is growing with exponential rates and thus it is a big challenge to collect effectively this medical data and then analyze and process for various insights [6]. These days, medical databases are getting data from sources like social media, mobile applications and digital marketing. All these sources makes Big Data Analytics a favorable choice for managing this healthcare data.

Government Regulations in Healthcare: Under the National eHealth Authority, India has decided to digitize the medical data in terms of Electronic Health Records (EHR) for the easy storage and exchange of medical data which makes its availability across all health institutions. Thus a technology like Big Data Analytics plays its prime role in storing and processing this data for ensuring better availability of health history of patients.

Organization in HealthCare Data:

Big Data is one of the best choice to extract structured data from Electronic Health Records which are unstructured in nature. Evidences out of observations can be extracted by applying Machine Learning over such Electronic Health Records without raising any kinds of issues.

Medical Knowledge Sharing: As lot of research is going on in medical sciences for improving healthcare systems, it becomes important to outreach this information among healthcare professionals for the access of such resources. With the help of Big Data Analytics, it is possible to provide interface for health professionals to update themselves with new medical information and thus to improve the healthcare delivery.

Early Diagnosis among Patients:

One of the big challenge in present healthcare systems is to take different opinions from various medical experts as every time it needs a lot of time to understand the previous patient history and other specific details. Big Data Analysis will turn to be handy for such medical experts to take decisions on the basis of insights provided by Big Data platforms.

5. Applications of Big Data Analytics in Health Care Systems:

There are number of applications in Healthcare systems where Big Data Analytics is playing an important role. Most important areas where Big Data finds its role are given below:

Electronic Health Records: It is a digital repository which stores information in terms of medical history, laboratory test reports and patient prescriptions. These records are shared via secure communication channels among private and public sector providers.

Real Time Alerting: Big Data Analytics is supporting Clinical Decision Support software which are providing on spot suggestions due to which the medical professionals can make their decisions.

Healthcare Predictive Analytics: Efficient Big Data pipelines are used for the detection and prediction of various diseases like Diabetes, Brain Tumors etc. Accurate detection models running on Big Data pipelines are making computations of data at lower costs and also providing results in less time intervals.

Medical Imaging: Medical Imaging is one of the area where Big Data Analytics is quite productive. The

analysis of Medical images is very expensive and time consuming process if done in manual ways. Big Data tools can effectively process medical images by using powerful algorithms for converting these images into numerical results which will help physicians for better diagnosis.

6. References

- [1] Kudyba, S. P. [2010]. Healthcare informatics: improving efficiency and productivity. CRC Press.
- [2] Mamoon Rashid, Mir Aamir, Balwant Ram [2018]. Big Data and Hadoop Computational Challenges - A Comparison Study. International Journal of Computer Sciences and Engineering, 6(7), pp.1446-1452.
- [3] Raghupathi, W., & Raghupathi, V. [2014]. Big data analytics in healthcare: promise and potential. *Health information science and systems*, 2(1),3.
- [4] Gui, H., Zheng, R., Ma, C., Fan, H., & Xu, L. [2016, November]. An architecture for healthcare big data management and analysis. In International Conference on Health Information Science (pp. 154-160). Springer, Cham.
- [5] Groves, P., Kayyali, B., Knott, D., & Kuiken, S. V. [2016]. The 'big data' revolution in healthcare: Accelerating value and innovation.
- [6] Feldman, B., Martin, E. M., & Skotnes, T. [2012]. Big data in healthcare hype and hope. *Dr. Bonnie*, 360, 122-125.

About the Authors



Mr. Mamoon Rashid is currently working as an Assistant Professor at Lovely Professional University, Jalandhar, India. His area of expertise is Big Data Analytics, Machine Learning and Neuroscience. He has published more than 20 International publications in Big Data Analytics and Machine Learning indexed in Scopus and Web of Science based Journals.



Dr. Harjeet Singh is presently working as an Associate Professor in the P.G. Department of Computer Science, Mata Gujri College, Fatehgarh Sahib, Punjab, India. He has published many papers in rated journals with indexing in Web of Science and Scopus and currently guiding several research scholars at Doctorate level.



Dr. Vishal Goyal is presently working as Professor in Department of Computer Science, Punjabi University, Patiala, India. He is Coordinator, Research Center for Technology Development for Differently Abled people and co-coordinator of Center for Artificial Intelligence and Data Science, Punjabi University, Patiala. His main research area is Cognitive Computing, Artificial Intelligence, Natural Language Processing and Machine Translation Language Technologies. He has



Mr. Kulvinder Singh is Final Year student in School of Computer Science & Engineering, Lovely Professional University, Jalandhar, India. His area of expertise is Big Data Analytics and Machine Learning.



Mr. Neeraj Mogla is currently working as Visual Insights & Analytics Manager in Nike Inc. Global Head Quarters, USA. He did his Master's in Computer Science from University of Akron, USA. Previously he worked as Senior Application Engineer for UnitedHealth Group. His main area of expertise is Data Analytics, Data Sciences, Machine Learning, Big data and Cloud computing.

Security Aspects in E-Commerce

► Gopal Krishna

Asst. Professor in CS & Engg. Dept. and
Student Branch Coordinator of CSI Netaji Subhas
Inst. of Technology, Patna, Bihar

► Aayush Aman

Secretary, Student Council, CSI Student Branch,
NSIT, Patna

► Tushar Anand

Student of B.Tech, 2nd year,
C.S.E Department of NSIT, Patna

Without proper trust, most reputed business operator and customers also may decide to get back to using traditional methods of business exchanges rather than the use of the internet for e-commerce activities. The threats to the network security at the e-commerce sites must be constantly looked at and measures should be implemented accordingly. The security measures should be implemented in such a way that they do not compromise with the intent of the website and its functionality. In this article, we will discuss the various persistent threats to network and computer security. Hackers as well as the e-commerce sites themselves are a reason for origination of such threats.

1. Introduction

The exchange or buying and selling of commodities using the internet on a large scale facilitating the feature of transportation from one place to another is known as e-commerce. However, the term e-commerce has a much wider spectrum and involves many more commercial activities than just shopping over the web. Some people and businesses use the terminology electronic business when talking about the wider aspect of the electronic commerce. In this paper, we will discuss the term e-commerce in its broadest definition and talk about the various security threats to which it is vulnerable. The trend of taking businesses online has seen a rapid growth over the past few years. The online transactions to take place in due course are made with the help of various forms. However, e-commerce has inevitably invited its share of problem creators. As much as e-commerce has been able to simplify transactions, it is oftenly plagued by serious threats that compromise its security as a medium for money and data exchange. This article is organized as: In section-2 various security threats are discussed, section-3 deals with the different measures to stay safe during online shopping and in section-4, conclusion of this article is mentioned.

2. Security Threats to E-Commerce

Security is the vital constituent which affects e-commerce which comprises of Computer Security, Data Security and other fields. Businesses

facilitating online payment systems are at higher risk of getting aimed at which can cause great losses to data and funds to the merchant. However, usually e-commerce systems provide adequate security systems to keep themselves safe. To educate the consumer is the critical part in terms of security. Some of the common yet treacherous methodologies behind the compromise in e-commerce security systems are discussed in below subsections.

2.1 DDoS

A DDoS or Distributed Denial of Service attack refers to sending an uncountable number of unwanted traffic to the server which results in overloading it. As a result, the server reaches to its superabundance state and start rejecting the legal connections.

Due to this attack the server might slow down otherwise causing it to totally shutdown becoming unresponsive.

According to a report, a huge DDoS attack was coordinated on 21st October, 2016 in which thousands of online shopping website were compromised causing a great loss. This attack was done with a series of phone calls, messaging and emails that resulted in condition of the server to get totally out of control. This attack is not about the credit card or debit card transactions, it's about the effect of those lost trusted buyers that would cripple your conversion rates and ultimately lead to the termination of a particular ecommerce website.

2.2 Clickjacking

Clickjacking, also termed as the

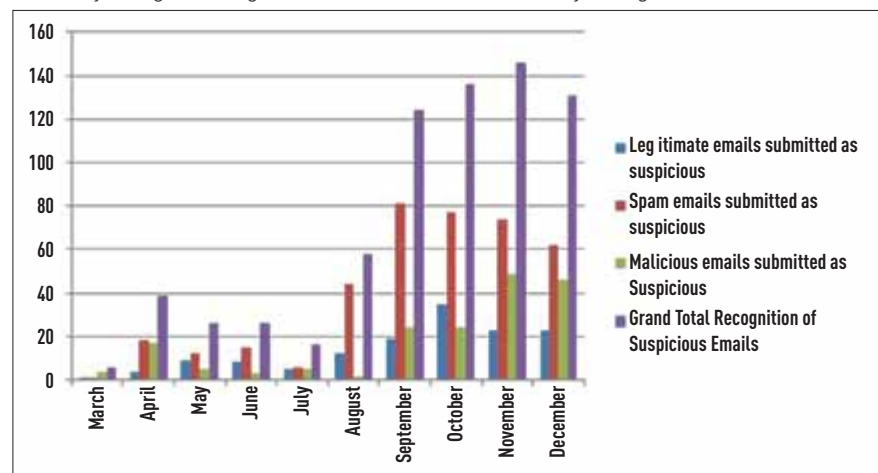


Fig 1: Recent Trends in Clickjacking (2018)

“UI redress attack”, is when a hacker creates multiple transparent or solid layers to attract the user to click a button and get redirect to another web page. Using this technique, the keystroke is copied too. Keystroke hijacking is used in e-commerce website with a full well-crafted combination of text, style sheets and image boxes which looks like original form and the user submit their respective details like username, password and bank account details. All the details are stored in attacker server and is later misused by the attacker for personal benefits of all sorts.

2.3 Ransomware

Ransomware is one of the modern threats to the e-commerce websites. It does exactly what the name suggests. Ransom Malware or as we simply say Ransomware is a type of malware that prevents admin from accessing their data from the website or server and then the attacker can threaten to compromise all the important data of the website servers within a limited time period until and unless one pays the ransom payment and then is able to regain access to the website.

To prevent data compromise from this threat, one must create the backup off all the information in regular time intervals and make sure the backup is totally up to date. If one has the updated data backup, after the exploitation the website can easily be restored formatting the locked-up data as well as the malware infused.



Fig 2: Ransomware

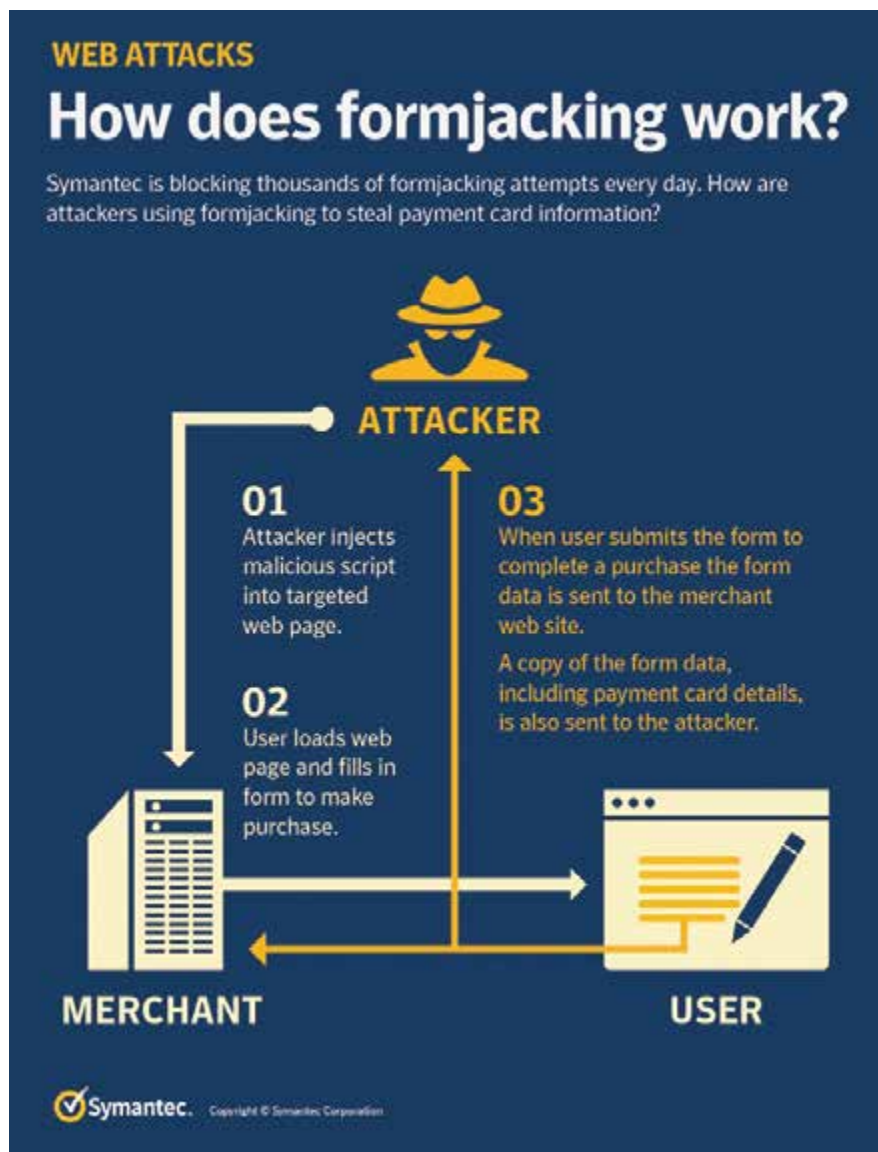


Fig 3: How does FormJacking Work?

2.4 XSHM

Cross-Site History Manipulation usually abbreviated as XSHM, is a SOP (Same Origin Policy) based security breach. Same origin policy is the most vital security concept used by the modern browsers. SOP states that web pages from different origin points cannot communicate with each other. Cross-Site History Manipulation technique is based on the fact that client-side browser history is not precisely partitioned on a per-site basis. Manipulation in the browser history can lead to compromise in the SOP, allowing the use of bi-directional CSRF (Cross-Site Request Forgery)

token and other exploitations such as the user privacy violation, login status detection, resources mapping, sensitive information inferring, user activity tracking and URL parameter stealing.

2.5 Trapdoors/Backdoors

While the development of the code of the e-commerce site, developers usually leave “trapdoors” or “backdoors” for the ease of monitoring the code throughout as it is developed. Instead of implementing a security protocol for accessing the code, backdoors route a quick way into the source code. Though it is convenient, trapdoors can lead to major security issues if they are not

taken care of prior to the launch of the e-commerce site. An attacker can easily use this vulnerability to exploit the site and cause system wide vandalism to the e-commerce server.

2.6 Formjacking

Formjacking attacks are quite simple to understand, it's a type of virtual ATM skimming - where cyber criminals try to inject some malicious code into the e-commerce websites to steal the customer's bank card details.

According to Symantec (An American based Cyber Security Company), about 4818 websites were exploited using formjacking malicious code every month in the year, 2018. According to the reports, in 2018, there were more than 3.7 million blocked attempts of formjacking by cyber security professionals. Some of the big companies compromised during the attack using formjacking methodology were Ticket Master, British Airways, Feedify, Newegg and many more to count.

3. Staying Safe Online- Security Measures

A companywide understanding of e-commerce security safety, techniques and dangers can enable every client and security administrators to believe the system that they're working with. At a comparable time, the large number of hacking and breaking applications accessible will bring about a genuine risk to e-commerce applications. Hence, it's quite required to comprehend security dangers beforehand, and notification

the most straightforward answers for minimize the dangers they face.

The understanding of e-commerce security safety techniques and threats on a wider level can enable every customer and security administrator to trust the system they are dealing with. If the right measures are taken to secure and formulate the framework, it's non-viable for an unauthorized client to gain access to the system and compromise it.

Some feasible solutions that can minimize the risk of getting compromised are:

- Use of Secure Socket Layer (SSL)
- Use of Public Key Infrastructure (PKI)
- Use of Premium Anti-Virus Software
- Security through Obscurity
- Educating Users

4. Conclusion

E-Commerce is usually understood as the purchase and selling of products online over the internet, but any transaction, either data or funds that is solely accomplished via electronic measures can be referred to as E-Commerce. The use of E-Commerce worldwide, has witnessed a rapid growth over the years, and still the people connected are increasing every day. E-Commerce has helped in the establishment of online retail marketing and has helped bloom businesses. E-Commerce security refers to the protection of assets from unauthorized access, use, fabrication, or destruction. Dimensions of E-Commerce security

include: Privacy, Authenticity, Integrity, Non-Repudiation, Confidentiality and Availability. Fraudsters are always on the look to compromise data taking advantage of the novice errors people tend to do. Common mistakes that leave customers vulnerable to security threats include the use of insecure websites, giving out more than required personal information and leaving computer systems open to viruses. In this paper, we discussed about the broader aspect of E-Commerce, dimensions of E-Commerce Security, E-Commerce Security threats (DDoS, Clickjacking, Ransomware, XSHM, Use of Backdoors and Formjacking) being some of the methodologies used to compromise security and guidelines for safe and secure online life being engaged in E-Commerce activities.

5. References:

- [1] <https://www.digitalattackmap.com/understanding-ddos/>
- [2] <https://www.cloudflare.com/en-in/learning/ddos/what-is-a-ddos-attack/>
- [3] [https://www.owasp.org/index.php/Cross_Site_History_Manipulation_\(XSHM\)](https://www.owasp.org/index.php/Cross_Site_History_Manipulation_(XSHM))
- [4] <https://www.symantec.com/blogs/threat-intelligence/formjacking-attacks-retailers>
- [5] <https://www.techgenyz.com/2018/12/18/e-commerce-security-threats-and-solutions/>
- [6] Poonam Patel, Kamaljeet I. Lakhtaria "A Study on E-Commerce Security Threats" *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 5, 2017.

About the Authors



Mr. Gopal Krishna [CSI-F8001991] has more than 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 learning, Wireless Sensor Networks and Online Social Networks.



Mr. Tushar Anand [CSI-01461042] Student of B.Tech, 2nd year, C.S.E department of NSIT, Patna. He is a tech geek and a knowledge seeker in the field of technologies related to android, Windows and various other platforms.



Mr. Aayush Aman [CSI- 01461040], Student, currently pursuing B.Tech (2nd Year) in C.S.E Department and currently positioned as Secretary, Student Council, CSI Student Branch, NSIT, Patna. He has a keen interest in pursuing his career as a Cyber Security Professional. He has coordinated in various CSI events.



CSI Region – IV Regional Student Convention

Organized by CSI Student Branch – GIET Gunupur in collaboration with CSI – Gunupur Chapter

24 – 25 February 2019

Hosted at : Department of CSE, GIET University



The CSI Region-IV Regional Student Convention was organized by CSI Student Branch – GIET Gunupur, in collaboration with CSI- Gunupur Chapter on 2019 on “Artificial Intelligence and Machine Learning” during 24-25 February 2019.

In this competitive scenario the aim of such an event was to bring out the best in students to nurture their talent to bring out their within capability and from them to produce the future masters of computer proficiency, artists, painters etc. Apart from the academic excellence, the motive behind the function is the overall development of the students. Educating and character building of youth calls for a high degree of responsibility and the prime concern is at achieving and grooming the interpersonal and the societal skills of the students.

Around 250 students from various parts of Region - IV actively participated in the event. Chief Guest Prof. (Dr.) Kishenkumar Tadisina Reddy, Vice Chancellor, GIET University, and Guest of Honour Prof. (Dr.) Senthil Kumar K, School of Engineering, GIET University showed their benign presence on the occasion of Inaugural ceremony.

This was followed by “Poster Presentation Competition” on the convention theme where the students participated with great zeal depicting their enormous talent. Different varieties of posters were presented by the participants on Information and communication technology. The session was chaired by Dr. Premansu Rath and Mr. Ranjeet Kumar Panigrahy.

This was followed by “Paper presentation” where various slides were shown regarding to this technology making the people think about it in and out in this technological advanced environment. In a time span of only 7 minutes students had generated interest and curiosity in the minds of the listeners. The session was chaired by Dr. Nilamadhab Padhy.

Technical Quiz competition and App Development competition was conducted and many teams from different Colleges/Institutes were participated. The session was judged by Er. K. Murali Gopal, Er. Rajiv Senapati and Er. D. Anil Kumar.



Day2 (25th Feb) started with thematic address by Prof. (Dr.) A. K. Nayak, during his speech he stated various examples to make the audience understand the ethical practices in information and communication world in a very effective manner.

After that “Project Presentation Competition” was organized where participants express and shared their views about the respective subject with so much of confidence and energy. Students linked the project with the presentation and used the presentation as means of communication showing various processes which involve Initiation, Planning, execution, and testing. The event was judged by Prof. (Dr.) A. K. Nayak and Prof. (Dr.) S. Kuannar.

“Valedictory Programme” conveyed the report of the two day convention presented by Prof. S. N. Das (SBC, CSI-SB GIET Gunupur), sending off message to the members and to also inspire students to embark on an stimulating quest for AI and ML.

Finally the day was ended with prize distribution by Prof. (Dr.) A. K. Nayak (Vice President, computer Society of India) and then followed by “Cultural events” in which participants enthusiastically displayed their talent by performing Group Song and Group Dance and enjoyed the evening.



Seminar conducted on 'D-Link Academy Campus Connect Program'

Organized by

Universal College of Engineering in association with CSI

Reported by **Mr. Pradeep Rath** Member, National ExecCom, RVP-VI, Maharashtra & Goa



(L to R) Mr. Sudhanshu Ojha of D-Link, Mr. Pradeep Rath RVP-6, CSI and Mr. V. Gosavi of IT Teamworks



Mr. Pradeep Rath RVP-6 CSI with Universal College CSI Student Body

'D-Link Academy Campus Connect - An Advantage of D-Link Global Certification to Student's Career' program was arranged at Universal College of Engineering in association with Computer Society of India. Program was enriched by the presence of CSI Western Region Director Dr. Pradeep Rath, D-link academy Head Mr. Sudhanshu Ojha and IT Team work manager Mr. Vijay Gosavi. D-link academy Head Mr. Sudhanshu Ojha delivered a session to Students on Advantage of D-Link Global Certification to student's Career. The seminar saw more than 60 attendees from second year and third year along with senior faculties Prof. Kanchan mam and Prof. Snehal mam followed with a Q&A session at the end of the session.

The program is based on concept of PBL-Project based learning which is demand for today's education system to match world standard. The Centre of Excellence Lab will address the Project Based Learning Program followed by global certification to increase the employability Index. D-Link Academy delivers a comprehensive, 21st century learning experience to help students develop the foundational ICT skills needed to design, build and manage networks, along with career skills such as problem solving, collaboration and critical thinking. Also the academy wish to work closely being DTE-Maharashtra as an Industry Partner to make UCOE campus a Digital Premises for better skill learning platform.

Dr. Pradeep Rath addressed students about the benefit of CSI membership and various new upgradation under CSI services. He also motivated students to actively participate in CSI internship and social awareness programs.

Mahatma Gandhi Mission's College of Engg. and Tech., Kamothe, Navi Mumbai

Workshop Report on Android Application Development

Organized by Department of Information Technology



(L to R) Dr. K. G. Narayankhedkar - Director General MGM facilitating Mr. Pradeep Rath, RVP-6 CSI in presence of Dr. S. K. Narayankhedkar, Principal MGM College

Introduction to Workshop

The Android Application Development Workshop was conducted by the IT Department of MGM on 5th & 6th April 2019. 36 participants participated in this Workshop.

A seminar was held on 5th of April 2019 on Android app development before the workshop begin. This seminar was held for emphasizing the importance of Android in today's era and encouraging the students to learn developing an Android applications.

Honorable DG Dr. K G Narayankhedkar and Respected Principal Dr. S K Narayankhedkar arrived with Chief Guest Mr. Pradeep Rath and Trainer Mr. Nikhil Pawar at 10.30 am.

After the lighting lamp, dignitaries were invited to be seated on the dias. Principal has addressed the participants CSI student to come out with inner potential without depending on the brand of college. He discussed about the new apps developed in the market. Director General has addressed to the students to develop the skill for improving their Knowledge. Also motivated the students to participate in the completion national and international level. Chief Guest Mr. Pradeep Rath has emphasized on soft skills, Entrepreneur skills and learn to take chances.

HOD Mr. Venkat K Raman introduced everyone and elaborated the Chief Guest Mr. Pradeep Rath's introduction. After chief guest, Convener of workshop Prof. Yogesh Sahare emphasized on importance of Android, day to day uses of it and importance of this workshop.

Day 1: Basic introduction of Android and the related software for the development of the application as the platform also covered the installation of Android, and pre requirement such as Java JDK version above 1.8, SDK and related compatibility technical things. How to create basic layout of the application which is extended by adding images, login page, email, text. How to call a toast and intent in an application.

Day 2: The session started with practicing the previous projects and each participant interact individually with the trainer. Some real project were developed as per the instruction, such as Android app to insert the audio app and video. How to get the permission for internet cameras and sensors. Participants were doing more hands on practices related this session.

More features of Android were introduced. During this session the participants were thought how to host the developed application in the play store. At the end of session, a photo session took place and many students addressed their views as a form of send off and thus the workshop was concluded successfully. At the end of the workshop students will get the following benefits.

- To induce the flavor of application development from the initial stages.
- Each of them would be in a position to develop their own application.
- The individual were expected to grasp the essence of the workshop thoroughly.



CSI Regional Student Convention- Region-III “KAUSHAL”-2019

Organized by Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore



CSI President Prof. A K Nayak lighting the lamp of CSI Regional Student Convention-Region-III



Delivering inauguration speech by Prof. A. K. Nayak, CSI President

The Nation-wide CSI-Conventions is continuously organized by different organizations. CSI Regional Convention is a platform for sharing ideas and expands networking for scholars, researchers, and students from all areas of Science and Technology.

The main aim of the Regional Student Convention is to build a foundation and inspire the young students of engineering to understand and apply the new trends in technology. The convention targets to give a broad overview of technical fusion. It invites students to participate in various events. As a Professional body, CSI encourages and supports R&D Projects in various branches of Engineering and Technology. It sponsors the Indian IT Professionals at the conferences and workshops.

This convention has 08 events which included Workshop, Research Paper & Articles, Project contest, Seminar, Mobile App Development, Programming Contest, Technical Quiz and Web Design Contest. The convention provides fruitful knowledge dissemination and cultural exchange among students from various parts of the country.

Introduction

A Regional Student Convention-Region III “KAUSHAL-19” was organized on 15th -16th March 2019 at Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore hosted by CSI, Indore Chapter and CSI-Student Branch, SVW Indore. Kaushal-19 is sponsored by M. P. Council of Science and Technology and Computer Society of India-Head Quarter. Convention theme was “Recent Digital Technology Advancement” is of great importance because of the digital technology application is transforming every walk of the life. It was our great honor and privilege to become organizer of Regional Student Convention Region-III. Computer Society of India Student Branch of Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore is one of the most active branch with large number of Student members and Activities.

Inaugural Session

Chairperson of Regional Student Convention-Region III “KAUSHAL-19”, Dr. Anand Rajavat gave brief introduction to

“KAUSHAL-19”. He detailed about all the contest and convention theme. Dr. Upinder Dhar, Vice Chancellor – Shri Vaishnav Vidyapeeth Vishwavidyalaya welcome Chief Guest, Sponsors, Judges and Participants. In his welcome address told about programs offered by Shri Vaishnav Institute of Information Technology with Industry collaboration like IBM, Redhat, NCSSS, Cisco and Apple with recent industry requirement on Big Data, Cloud & Mobile Computing, Enterprises System, Cyber Security and Information System, Artificial Intelligence. He share the objective of this Regional Convention is to provide a phenomenal opportunity for the students and importance of platform of Regional Student Convention. Chief Guest of Regional Student Convention-Region III “KAUSHAL-19” Dr. A.K. Nayak, National Vice President cum President (Elect) of Computer Society of India highlighted the need for Recent Digital Technology Advancement in Modern era. The society needs to build an innovative ecosystem by combining the use of this advanced technology with the best minds across Academia, Business, Industry and Government for the overall benefits of the common man. He told on futuristic research are in Machine Learning, Artificial Intelligence, Data Science, IOT & IOE. He encouraged students to focus on all the area of possibility to become a Researcher scientist and always being positive to become a successful person. At the end, the words of thanks were proposed by Mr. Vijay Prakash, State Student Coordinator (M.P.) and CSI-SB, SVW Branch Coordinator.

Valedictory Session

On Second day of Regional Student Convention-Region III “KAUSHAL-19” In Valedictory Session, Chief Guest Mr. Sanjeev Agrawal, Vice President-Operations and Human Empowerment, Impetus Infotech (India) Pvt. Ltd., Indore Dr. Upinder Dhar, Vice Chancellor – Shri Vaishnav Vidyapeeth Vishwavidyalaya, Dr. Anand Rajavat, Coordinator- Shri Vaishnav Institute of Information Technology, Mr. Vijay Prakash, Convener of Regional Student Convention-Region III “KAUSHAL-19” were present on dais. In the speech Dr. Upinder Dhar said that in this period of change innovation is very important and congratulate to participants of different

activity and motivate to implement their idea in large level, Chief Guest Mr. Sanjeev Agrawal said that He appreciated the event "KAUSHAL-19" to provide platform to Students for showing their Ideas, Concepts & Thoughts. Mr. Vijay Prakash gives brief report of "KAUSHAL-19", according to that total 500 students from different Institute across India participated during two days Convention "KAUSHAL-19". Mr. Vijay Prakash, Convener of Regional Student Convention-Region III "KAUSHAL-19" informed that cash prizes of a total Fifty thousand rupees was distributed to the winners and also to runner up teams as well as total Fifty thousand rupees in form of vouchers & coupons was distributed to the winners and also to runner up teams. At the end, the words of thanks were proposed by Dr. Anand Rajavat, Chairperson – KAUSHAL 2019.

Technical Seminar

On First day of Convention, Seminars on Recent Digital Technology Advancement given by TCS Expert Mr. Ashish Gangele. He Conducted a session on Java Web Application Frameworks, Application Architecture, Single Page Application in Java Phase 1 - AngularJs and Phase 2- Angular 5 with Nodejs. DevOps and CICD. Approx 200 Students were attended the technical Seminar.

Research Publications & Articles

On First day of Convention, total 32 Research Papers and Articles were submitted. In this event, participants presented Research Publications & Articles.

Senior Professors and industrialist from renowned organization reviewed the research papers. The Judges for Research Papers and Articles were

- Dr. Vivek Uprit, Head of the Department, SOCA, SIMS Indore.
- Mr. Ashish Agrawal, Senior Software Engineer, Systematin Infotech, Indore

Programming Contest

On First day of Convention, total 45 Participants were participated in Programming Contest. In this contest, the participant had to find out various bugs in the given code and had to remove them all. Program statement allotted on the spot to the participants for which the participant had written code for a given problem. Mr. Mayank K. Sharma, CEO & Cofounder, AMSTECH Pvt. Ltd. was the Judge for Programming Contest.

Technical Quiz

On First day of Convention, total 55 Participants were participated in Technical Quiz. The quiz were restricted in information and communication technology, and computer

science discipline. This event primarily focused on the participant's ability to give correct answers of all the given questions during the limited period of time. Mr. Anurag Agrawal Founder Jawabdo.co.in was the Quiz Master.

Project Contest

On First day of Convention, total 44 Participants were participated in Project Contest. In this event UG/PG students had to present their project along with running code. The project presentation must primarily focused on Innovative approach, Technical expertise, Benefits for the development of society, Applications and real-time wide usage of the project.

Senior Professors and industrialist from renowned organization Judge the participant Projects. The Judges for Project Contest were

- Mr. Jitendra Patidar from FCC Kuwait (Towx Infotech Pvt. Ltd.)
- Mr. Suryakant Tripathi from Medicaps University

Technical Workshop

On First day of Convention, total 20 Participants were participated in Technical Workshop. Workshops related to recent Digital Technology Advancements in the field of IT. Workshop topic focused on Internet of Thing. The workshop conducted in association with IBM and ORACLE. Resource Person of Workshop was Mr. Deepak Shrivastava, Technical lead from WebTek Pvt. Ltd. Delhi, IBM & Oracle Authorized Partner.

Web Design and Mobile Application Development Contest

On Second day of Convention, total 28 were participated in Web Design and Mobile Application Development Contest. The Web Designing Contest based on all about imagination, adding interactivity to web pages, Applying responsive design to enable the page to be viewed by various devices, describe the basics of Cascading Style Sheets (CSS3) and Use the Document Object Model (DOM) to modify pages. Engineers usually see the world as an individual with a curious mindset. This contest made students explores the vast areas of creativity and Innovative ideas in WEB DESIGNING and DEVELOPMENT.

Senior Professors and industrialist from renowned organization evaluated Application. The Judges for Web Design and Mobile Application Development Contest were

- Mr. Devendra Mistri, Production Engineer, Digilet Indore
- Mr. Jitendra Jain, Technical Lead, Value Labs.

More than 200 Participants from various institutes across India in Regional Student Convention-Region III "KAUSHAL 2019" including Indore, Gwalior, Bhopal, Chindwara, Khargone as well as from Rajasthan, Gujarat and Chhattisgarh had participated. ■





Computer Society of India

<http://www.csi-india.org>

Call for Regional & State CSI Student Coordinator/s - 2019-2020

Computer Society of India, with 500+ Student Branches across India with close to one lakh members invites applications from Life Members of CSI interested in serving as Regional and State Student coordinators. The Student Coordinators should be able to contribute to the growth of CSI Student Branches and to increasing activity towards spreading of IT Education in different academic institutions.

The National Student Coordinator (NSC) / Regional Student Coordinator/s (RSCs) will work closely with the Execcom, Regional Vice Presidents and CSI Education Directorate for leading, coordinating and managing student activities in the seven regions in the country. RSCs will be supported by the State Student Coordinators (SSCs) in each of the states which constitute the Regions. Details of the Regions and constituent states are given below in this advertisement.

The applicants must be dynamic with experience of serving the student community and with ability to organize Student Conventions. CSI has ambitious plans of growth and intends to conduct programs in line with the national IT policy of the Government of India so as to prepare the student community to contribute towards nation building.

The student coordinators at the Regional and State level must be Life Members of CSI with experience of serving CSI at the Chapter level and should be ready to devote time to carry out the tasks for the benefit of CSI Student Branches. Preference will be given to applicants who have organized Institutional / State level / Regional / National events and are working in an academic institution.

Interested candidates are requested to send their detailed CV (including Qualification, Experience, Position held at CSI & Membership Number) by email to adm.officer@csi-india.org with copy to Hon Secretary at secretary@csi-india.org / drskyadav@hotmail.com and Prof P Kumar, National Student Coordinator at pkumar_5@yahoo.com mentioning details of the position applied for on or before 5th June, 2019.

The following is the list of Regions and States for which Student Coordinators are invited:

Regions	States
Region-1	Delhi
	Haryana
	Himachal Pradesh
	Jammu & Kashmir
	Punjab & Chandigarh
	Uttar Pradesh
	Uttarakhand
Region-2	Assam
	Bihar
	West Bengal
Region-3	Gujarat
	Madhya Pradesh
	Rajasthan
Region-4	Chhattisgarh
	Jharkhand
	Odisha
Region-5	Andhra Pradesh
	Telangana
	Karnataka
Region-6	Goa
	Maharashtra
Region-7	Kerala
	Tamil Nadu 1 & Puducherry
	Tamil Nadu 2

Prof. A K Nayak
President, CSI

Dr. Santosh Kumar Yadav
Hon. Secretary, CSI

STUDENT BRANCHES INAUGURATION

REPORTS

University of Engineering & Management (UEM), Kolkata, West Bengal, (Region-II)



University of Engg. & Mgmt. (UEM), Kolkata opened the student branch of CSI on 13th April 2019 at UEM Kolkata campus. The inaugural ceremony was graced by the presence of CSI Kolkata chapter leadership team comprising Mr. Gautam Hajra, Chairman, Mr. Aniruddha Nag, Vice Chairman, Mr. Sourav Chakraborty, Secretary and Mr. Snehasis Banerjee, Treasurer. The program was enriched by the presence of distinguished speaker of the day Professor Dr. Amlan Chakraborty, Dean CU. Prof. Sukalyan Goswami, Registrar and HOD, CS & Engg. (CSE), UEM Kolkata welcomed the distinguished guests and addressed the audience with his welcome note emphasizing the significance on establishing the branch. Distinguished CSI Leaders explained in detail the benefits students can obtain in the field of Industry, research and academia by joining branch. They also explained the students about the importance of various Cutting Edge Technologies in the field of CS & Engg., and how CSI can play a pivotal role in development of such skills among the CSI student members. They also shared their experiences about the changes taking place at IT industry in the direction of automation, cloud computing, robotics, machine learning, AI, IoT and so on. The program was enlightened by Prof. Dr. Amlan Chakraborty, Dean CU, with his vibrant talk session on IoT. The interactive session raised lot of enthusiasm among the students regarding prospect of IoT and the implementation in human life. Students have shown lot of interests about CSI and the technical discussion taken place during the day. Prof. Kaustuv Bhattacharjee, UEM Kolkata thanked all the speakers of the day with an expectation of venturing into various knowledge intensive activities in future by organizing seminars, workshops, participation in conferences under the guidance of CSI.



Dnyanshree Institute of Engineering & Technology, Satara, Maharashtra (Region-VI)



RWMCT's Dnyanshree Institute of Engineering & Technology, Satara, Maharashtra has inaugurated CSI Student Branch on 19th March 2019 under the guidance of Mr. D B Wangde, Hon Chairman. Mr. Rohit D Wangde, Managing Director Inaugurated the Student Branch. At this auspicious occasion, Prof. Dr. Ajay D Jadhav Principal, Mr. Mohan Kulkarni, Prof. O C Nilakhe, Prof. Pradip N Shendage [CSI Faculty Coordinator], Prof. Homkar, Prof. Dhane, Prof. Pondkule, Prof. Padwal & Prof. Gurav were present. The event was witnessed by around 100 delegates including CSI members, faculty members and students. Hon Managing Director Mr. Rohit D Wangde was inaugural keynote speaker. He congratulated CSI membership. Prof. Dr. Ajay D Jadhav marked the importance of CSI role through the membership to the students, many of which belong to rural background. Mr. Nilakhe, stressed upon the significance, benefits and role of this branch for organizing professional activities. The branch will play an important role in achieving the objective of the CSI. CSI Student branch has proposed to have connectivity to the professional IT environment. To enhance the knowledge of members, the branch will take efforts to organize different workshops and conferences, guest lectures, poster presentation / exhibitions, seminars. It will stimulate the research culture in the faculty and students. The program concluded with vote of thanks.



CHENNAI CHAPTER



CSI Chennai Chapter along with other associations hosted a technical talk on Opportunities in Open Source by Mr. Shakthi Kannan, Senior DevOps Engineer, Aerospike Inc, Bangalore, on Saturday, 2nd March 2019. The "Opportunities in Open Source" talk is an interactive session for the benefit of both students and IT professionals to know the possibilities that exist when working with Free (Libre) and Open Source Software projects. The Five Ws and How approach will be used to elaborate the necessary steps that participants need to adopt for their prospective future and career growth.



Chennai Chapter along with other associations hosted a technical talk on "Product Nation: India's Trillion Dollar Opportunity in SaaS" by Mr. Suresh Sambandam, Founder & CEO, OrangeScape on Friday, 29th March 2019. Over the last 30-40 years, due to the IT services boom, our country has built deep-domain expertise in basically every industry vertical. It is that expertise that needs to be productized to make India a global SaaS hub. Suresh gave the highlight on India's natural strengths in SaaS, and what it will take to make India a product nation and a Global SaaS hub. Chennai Chapter organised CSI Day Celebrations on 15th March 2019. Dr. S Poonkuzhali, Chairperson gave the Welcome address. The Celebrations were started with

project presentation with demo by the shortlisted 10 teams. MC members Dr. B Swaminathan, Dr. U Karthikeyan and Dr. Prema Kirubakaran were the Jury members of this event. First prize is bagged by SSS Shasun Jain College for Women, Second Prize by Rajalakshmi Engineering College and Third Prize by Panimalar Engineering College. It was followed by ICT Quiz by Mr. H R Mohan, Fellow and Past President, CSI. Vote of thanks is given by Mr. Anantha Padmanaban, MC Member.



KANCHEEPURAM CHAPTER



CSI Kancheepuram Chapter organized a technical talk on Data Lake Security on 5th April 2019 at Hindustan institute of Technology and Science, Padur, Chennai. Dr. K M Mehata, Vice Chairman and Dr. M Senthil Kumar, Hon Secretary of CSI Kancheepuram Chapter were presided over the event. The program starts with welcome address by Dr. Rajeswari Mukesh, Head, School of Computing Sciences, HITS. The Resource Person Mr. M Senthil Kumar, Project Director & Senior Consultant has been introduced by Dr. S Sathyalakshmi, Professor, HITS. The Resource person started with the topic of Data Lake Security. He has shared his knowledge on data lake security, data value chain, data warehouse and data lake etc. It was a very interesting and interactive with all faculty members. He has explained about the future technologies such as Artificial Intelligence, Big Data Analytics and Data lake, which will give more opportunity in future. He interacted with faculty members about concepts of data lake security. More than 60 CSI members have attended the session. The vote of thanks has given by Dr. M Senthil Kumar, Hon Secretary CSI Kancheepuram Chapter.

PUNE CHAPTER



The CSI Regional Project Competition for Region VI, organized by CSI Pune Chapter and Dr. D Y Patil Institute of Technology, Pimpri, Pune on 4th April 2019. Projects are the inevitable part of the Computer Engineering and Information Technology. The project competitions like this provide a great platform to the students to present technical skills, innovative ideas and soft skills. The competition received an overwhelming response with 71 entries from different colleges from Nasik, Kolhapur and Pune cities. All these projects were evaluated by the Panel of judges from Industry and Academics. 49 projects out of 71 projects were shortlisted for the on-the-Spot evaluation. The Event started with the registration, refreshments, an informal

inauguration and a common address to all participants. Prof. Smita Chaudhari, Student Branch Coordinator, has given a brief idea about the CSI Project Competition, the nature of the regional round. The event was inaugurated by the Chief Guest, Mr. Abhay Pendse, Chairman, CSI, Pune Chapter. Sandeep Bedi, was the Guest of Honor for the function. Icertis, Pune and Softcorner, Pune has sponsored this event. The event was coordinated by Prof. Smita Chaudhary, Prof. Pradnya Kulkarni and Prof. Rahul Pitale.

The winners of the competition:

Kevin Daas, VR Space Explorer from Sinhgad Institute of Technology and Science, Narhe, Pune.

1st Runner Up: Abhishek Gaikwad, Super Resolution using Deep Learning from Pimpri Chinchwad College of Engineering, Pune.

2nd Runner Up: Sneha Patil, Recommendation System for Precise Farming from Pimpri Chinchwad College of Engineering, Pune.

Apart from this, there were 3 special awards:

Best Social Aspect Project : Akshay Jumale, Pune Institute of Computer Technology.

Best Innovation Project: Vedant Patil, Dr. D.Y. Patil School of Engineering and Technology, Pune.

Dr. J.S. Umale Futuristic Scholarship Award: Rajeshwari Rajput, Dr. D.Y. Patil Institute of Technology, Pimpri, Pune.

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 Dr. R R Deshmukh, Chief Editor in the email id: rrdeshmukh.csit@bamu.ac.in with a copy to Prof. A K Nayak, Publisher, in the email id : aknayak@iibm.in and Dr. Brojo Kishore Mishra in email id: brojomishra@gmail.com.

Prof. A K Nayak
Publisher

FROM CSI STUDENT BRANCHES

REGION-I

Vidya College of Engineering, Meerut



9-3-2019 - Event on Technotsav (A CSI Technical Fest)

Guru Nanak Dev Engineering college, Ludhiana



26-2-2019 to 1-3-2019 & 5-3-2019 to 6-3-2019 Workshop on Python

REGION-III

Sarvajanik College of Engineering & Technology, Surat



19-3-2019 - Expert talk on iOS Programming
Dharmsinh Desai University, Nadiad

Gyan Ganga Institute of Technology and Sciences, Jabalpur



30-3-2019 - Workshop on Java

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



7-4-2019 - Workshop on Web Designing



5-3-2019 to 9-3-2019 - STTP on Internet of Things (IOT) : Enabling Technologies, Applications, Challenges & Research Opportunities

REGION-V

Chalapathi Institute of Engineering and Technology, Guntur



30-3-2019 to 31-3-2019 - International Conference on Recent Advances in Computer Science & Engineering

Malla Reddy College of Engineering, Secunderabad



16-2-2019 - Event on Importance & Benefits of CSI Membership by Dr. N J Rajaram

REGION-V

Anurag Group of Institutions, Hyderabad



22-3-2019 – Workshop on Design Thinking



22-3-2019 & 23-3-2019 – Event on HACK-PY (Hackathon on Python)

G Pullaiah College of Engineering & Technology, Kurnool



3-4-2019 – Mini Project Expo

JSS Academy of Technical Education, Bangalore



10-4-2019 & 11-4-2019 – 5th National conference on Advancements in Information Technology (NCAIT-2019)

GSSS Institute of Engineering & Technology for Women, Mysore



23-3-2019 – Hands-on Session on Design and Analysis of Algorithms



29-3-2019 & 30-3-2019 – Workshop on Building Internet of Things Applications using Raspberry Pi & Arduino

Maharaja Institute of Technology, Mysore



12-4-2019 – Technical Talk on IOT & its applications



20-2-2019 – Technical Talk on Recent Trends in Data Analytics

REGION-V

B.M.S. Institute of Technology & Management, Bangalore



22-3-2019 & 23-3-2019 – Workshop on Network Simulation using NS2

S.G. Balekundri Institute of Technology, Belgaum



4-4-2019 - Technical talk on Cyber Law

New Horizon College of Engineering, Bangalore



22-3-2019 & 23-3-2019 - State Level Intercollegiate Technical Festival and Competition



5-4- 2019 - National Level Student Technical Symposium – Yukthi 2K19

K S Institute of Technology, Bangalore



27-2-2019 - Seminar on Robotics and Home Automation



7-3-2019 - Computer Awareness Program

KLE Institute of Technology, Hubballi



12-1-2019 – Workshop on Professional Ethics



16-3-2019 - Skill Development Programme on Data Science

REGION-VI

MIT Academy of Engineering, Pune



29-3-2019 & 30-3-2019 - National Conference on Emerging Trends in Computer Engineering and Technology



1-4-2019 - Project Expo 2019 : National Level Project Competition

REGION-VI

Tatyasaheb Kore Inst. of Engg. & Tech., Warananagar



18-3-2019 20-3-2019 - FDP on Machine Learning using Python

REGION-VII

Ramco Institute of Technology, Rajapalayam



8-3-2019 - Real-time Coding Workshop

REGION-VII

S A Engineering College, Chennai



8-3-2019 - Mr H R Mohan, Fellow and Past President during CSI Foundation Day Celebrations

Sathyabama Institute of Science and Technology, Chennai



3-4-2019 to 5-4-2019 - International Conference on Artificial Intelligence and Machine Learning (IAIM 2019)

National Engineering College, Kovilpatti



28-2-2019 & 1-3-2019 - National level Technical Symposium NECSI'19



Student branches are requested to send their report to sb-activities@csi-india.org

Chapters are requested to send their activity report to chapter-activities@csi-india.org

Kindly send **High Resolution Photograph** with the report.

पधारो महारा देश जी



27 - 29 September, 2019



**53RD ANNUAL
CONVENTION**

Hosted by :

Computer Society of India, Udaipur Chapter

www.csi-2018.org

Starts 27th September, 2019

Become a part of 53rd Annual Convention, Get in Touch...

**For Sponsorship &
Exhibition Queries**

Mr. Gaurav Kumawat
Convener, 53rd Annual Convention
Mobile - +91 86969 32741

**For Accommodation &
Hospitality Queries**

Mr. Dinesh Sukhwal
Hon. Secretary, CSI Udaipur Chapter
Mobile - +91 94143 17017

**For Registration &
Transportation Queries**

Dr. Bharat Singh Deora
Chairman, CSI Udaipur Chapter
Mobile - +91 98295 88286

**For General Queries
& Support**

Mr. Amit Joshi - Immd. Past Chairman,
CSI Udaipur Chapter
Mobile - +91 99046 32888

csiudr@gmail.com