www.ijecs.in

International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 3 Issue 12 December 2014, Page No. 9430-9436

Web Intelligence-An Emerging vertical of Artificial Intelligence

Rahul Pareek

(BCA, MCA, MCTS)

Sinhgad Institute of Management, vadgaon (bk), Pune, India

rahulpareek2440@gmail.com

ABSTRACT: The information footprints of a rapidly increasing influx of internet users present us with an immense source of information that ultimately contributes to the construction of innovative web technology suitable for the future generations. Likewise, WEB INTELLIGENCE has been presented as the usage of advanced techniques in ARTIFICIAL INTELLIGENCE and INFORMATION TECHNOLOGY for the purpose of exploring, analyzing, and extracting knowledge from web data. Web Intelligence is the particular field of research and development to explore the fundamental roles as well as practical applications of artificial intelligence which is slowly developing throughout the world in order to create next generation of products, services and frameworks based on internet.

This paper outlines important aspects of Web Intelligence (WI), in the context of Artificial Intelligence in Education (AIED) research, research from the viewpoint of Brain Informatics (BI), Web Intelligence Technology and its Applications, Advances in Web Intelligence and Future & Trends of Web Intelligence in Artificial Intelligence.

KEYWORDS: Web Intelligence, Artificial Intelligence, Web Farming, Web Mining, Web Spiders.

I. INTRODUCTION:

WEB INTELLIGENCE: The term "Web Intelligence (WI)" was first introduced in 2000. As a new field of study, it presents excellent opportunities and challenges for the research and development of new generations of Web-based information processing technology, as well as for exploiting Web-based advanced applications. WI may be viewed as applying results from existing disciplines (e.g., Artificial Intelligence (AI) and Information Technology (IT)) to a totally new domain - the World Wide Web (the Web for short); WI may be considered as an enhancement or an extension of AI and IT; WI introduces new problems and challenges to the established disciplines. WI has been recognized gradually as a new research field on studying intelligence on the Web and intelligence for the Web. Web Intelligence is a multidisciplinary area dealing with utilizing data and services over the Web, to create new data and services using Information and Communication Technologies (ICT) and intelligent techniques.

ARTIFICIAL INTELLIGENCE: Artificial intelligence is the branch of computer science concerned with making computers behave like humans. The term was coined in 1956 by John McCarthy at the Massachusetts Institute of Technology. Artificial intelligence includes the following areas of specialization:

- Games Playing: programming computers to play games against human opponents.
- Expert systems: programming computers to make decisions in real-life situations (for example, some expert systems help doctors diagnose diseases based on symptoms).

- Natural Language: programming computers to understand natural human languages.

- Neural networks: Systems that simulate intelligence by attempting to reproduce the types of physical connections that occur in animal brains.

- Robotics: programming computers to see and hear and react to other sensory stimuli.

II. WEB INTELLIGENCE & ARTIFICIAL INTELLIGENCE IN EDUCATION:

WI explores the fundamental roles as well as practical impacts of Artificial Intelligence (AI) and advanced Information Technology (IT) on the next generation of Web-related products, systems, services, and activities. As a direction for scientific research and development, WI can be extremely beneficial for the field of AIED. Some of the key components of WI have already attracted AIED researchers for quite some time – ontologies, adaptivity and personalization, and agents. It focuses more on other issues in WI, such as intelligent Web services, semantic markup, and Web mining, and proposes how to use them as the basis for tackling new and challenging research problems in AIED.

The scope of WI as a research field, as proposed by Zhong et al. (2002), encompasses Web information systems environments and foundations, ontological engineering, human-media interaction, Web information management, Web information retrieval, Web agents, Web mining and farming, and emerging Web-based applications. It also aims at deepening the understanding of computational, logical, cognitive, physical, and social foundations as well as the enabling technologies for developing and applying Web-based intelligence and autonomous agents systems (Liu et al., 2003).

We can study Web intelligence on at least four conceptual levels (Zhong et al., 2002):

- Network level: Internet-level communication, infrastructure, and security protocols, where intelligence comes from the Web adaptivity to the user's surfing process.
- Interface level: intelligent human-Internet interaction, e.g. personalized multimedia representation.
- Knowledge level: representing (in machine-understandable formats) and processing the semantics of Web data.
- Social level: studying social interactions and behavior of Web users and finding user communities and interaction patterns.

SEMANTIC WEB: An important part of the background and context for discussing the relationship between WI and AIED is the Semantic Web. It is the new-generation Web that makes possible to express information in a precise, machine-interpretable form, ready for software agents to process, share, and reuse it, as well as to understand what the terms describing the data mean. It enables Web-based applications to interoperate both on the syntactic and semantic level (Hendler, 2001).

Key components of the Semantic Web technology are (Preece and Decker, 2002):

I. a unifying data model; currently, RDF(Resource Description Framework) is most frequently used data model on the Semantic Web.
II. languages based on RDF, such as DAML+OIL (DARPA Agent Markup Language plus Ontology Inference Layer), http://www.daml.org/2001/03/daml+oil-index, for developing ontologies and for marking up Web resources; semantically annotated Web resources, in turn, enable semantically rich service-level descriptions (such as DAML-S, the DAML-based Web Service Ontology, http://www.daml.org/services/).
III. Ontologies of standardized terminology to represent domain theories; they enable construction of support tools that assist the generation and processing of semantic markup of Web resources.

LANGUAGES: There are a lot of languages for developing ontologies and semantically annotating Web pages. One way or another, most of them are based on XML (eXtensible Markup Language), XML Schemas, RDF (Resource Description Framework), and RDF Schemas, all four developed under the auspices of the World-Wide Web Consortium (W3C) and using XML syntax (Gómez-Pérez and Corcho, 2002). Another important branch of languages is that for supporting WI-infrastructure issues, such as WSDL (Web Services Description Language), WSFL (Web Services Flow Language), UDDI (Universal Description, Discovery, and Integration), SOAP (Simple Object Access Protocol), and PSML (Problem Solver Markup Language) – see (Zhong et al., 2002; Liu et al., 2003; Preece and Decker, 2002) for starting points on the use of these languages to support development of WI.

III. WEB INTELLIGENCE RELATED WORK IN AIED:

AIED researchers study such issues in the context of teaching and learning theories and systems. For example, there is an extensive research and development effort in pedagogical agents, autonomous software entities aimed at supporting human learning by interacting with students/learners and authors/teachers and by collaborating with other similar agents, in the context of interactive learning environments (Johnson et al., 2000). Pedagogical agents can help very much in locating, browsing, selecting, arranging, integrating, and otherwise using educational material on the Web.

The work in Web-based intelligent tutoring systems (ITS) also has a long tradition. Web-based ITS are generally important for WI since they demonstrate how different intelligent techniques can be deployed to support a number of issues highly relevant for the learning and teaching processes on the Web, such as personalization, adaptivity, and collaboration, to name but a few. Although in Web-based ITS all such issues are learner-centered, their importance overcomes the domain of education. With some effort, parts of the Web-based ITS technology can be transferred to other application domains.

First-wave Web-based ITS like ELM-ART (Brusilovsky et al., 1996) and PAT Online (Ritter, 1997), to name but a few, were followed by a number of other learning environments that used Web technology as means of delivering instruction. More recent Web-based ITS address other important issues, such as integration with standalone, external, domain-service Web systems (Melis et al., 2001), using standards and practices from international standardization bodies in designing Web-based learning environments (Retalis and Avgeriou, 2002), and architectural design of systems for Web-based teaching and learning (Alpert et al., 1999), (Mitrović and Hausler, 2000). Rebai and de la Passardiere try to capture educational metadata for Web-based learning nvironments (Rebai and de la Passardiere, 2002).

A rapidly growing branch of AIED research is teaching and learning ontologies and ontology-aware authoring tools. Generally, ontologies provide the necessary armature around which knowledge bases should be built, and set grounds for developing reusable Web-contents, Web-services, and applications (Devedžić, 2002). The most notable classical work in the AIED community related to the development of educational ontologies comes from the Mizoguchi Lab at Osaka University, Japan (e.g., see (Mizoguchi and Bourdeau, 2000)), and from Tom Murray (1998).

IV. WEB INTELLIGENCE AND PERSONALIZATION OF LEARNING:

Adaptivity of Web-based systems plays an important role in WI (Liu et al., 2003). Important issues related to adaptivity of Web-based learning environments, such as providing adaptive navigation support to the learner, links annotations, and adaptive curriculum sequencing, have been already studied in the AI community (Brusilovsky, 1999). In the setting from Figure 2, an essential aspect of educational servers' adaptivity is personalization – they should be able to personalize interactions with each learner by keeping track of his recent visits/activities and relating the topics he learns and the sites he accesses during different learning sessions. Moreover, an intelligent educational server should actively help the learner and interact with him when executing these tasks. Since educational servers are interconnected, a specific server may personalize the session with a particular learner by pre-fetching the material the learner needs from other servers. This is an adaptive process based on observations of the learner's surfing behavior.It belongs to network level WI, and is different from interface-level WI, which is related to adaptive cross-language processing, personalized multimedia representation, and multimodal data-processing capabilities (Zhong et al., 2002).

V. WEB INTELLIGENCE MEETS BRAIN INFORMATICS:

There are urgent needs and great benefits of combining WI and BI research. Fundamental issues in both fields need to be investigated and integrated systematically in order to materialize those benefits.

Brain Informatics (BI) is an emerging interdisciplinary field to study human information processing mechanism systematically from both macro and micro points of view by cooperatively using experimental, computational, cognitive neuroscience and advanced WI centric information technology. It attempts to understand human intelligence in depth, towards a holistic view at a long-term, global vision to understand the principles and mechanisms of human information processing system (HIPS), with respect to functions from perception to thinking, such as multi-perception, attention, memory, language, computation, heuristic search, reasoning, planning, decision-making, problem-solving, learning, discovery and creativity. BI can be regarded as brain science in WI centric IT age [98, 99]. BI is proposing to study human brain from the viewpoint of informatics (i.e., human brain is an information processing system) and use informatics (i.e., WI centric information technology) to support brain science study. In order to study BI systematically and give a global view to answer what is brain informatics, we list several major subtopics in each research area below, which is an extensional description of BI research.

– Thinking centric investigation of HIPS:

•Human deductive/inductive reasoning mechanism for understanding the Principle of human reasoning and problem solving.

•Human learning mechanism for acquiring personalized student models in an interactive learning process dynamically and naturally.

– Perception centric investigation of HIPS:

• Human multi-perception mechanism.

• Auditory, visual and tactile information processing.

– Modeling human brain information processing mechanism:

• Neuro-mechanism of HIPS.

•Mathematical models of HIPS;

• Cognitive and computational models of HIPS.

– Information technologies for management and use of human brain data.
• Human brain data collection, pre-processing, management, and analysis.
• Multi-media human brain data mining and reasoning.
• Data basing the brain and constructing data brain models.
• Developing brain data grid and brain research support portals.

VI. "WI MEETS BI" IN PRINCIPLE:

As pointed out by McCarthy, if we understood enough about how the human intellect works, we could simulate it. However, so far we did not have sufficient ability to observe ourselves or others to understand directly how our intellects work. Understanding the human brain well enough to imitate its function requires experimental and theoretical success in cognitive science and neuroscience. Neuroscience, the study of the brain and nervous system, is beginning to have direct measurement and observation of ourselves or others to understand
directly how our intellects work. These measurements and observations are, in turn, challenging our understanding of the relation between mind and action, leading to new theoretical constructs and calling old ones into question. New instrumentation (fMRI etc.) and advanced information technologies are causing an impending revolution in WI and brain sciences. This revolution is bi-directional:

– WI for BI: WI based technologies (e.g., the wisdom Web, data mining, Multi-agent and data/knowledge grids) will provide a new powerful platform for brain sciences.

– BI for WI: New understanding and discovery of the human intelligence models in brain sciences (e.g., cognitive science, neuroscience, and brain informatics) will yield new WI research and development.

Importance of a Conceptual Model of the Brain:
It is a common practice to compare the human and a computer in order to gain an understanding of the one with the aid of the other. For example, one may draw many correspondences between them, including CPU to the brain, input/output devices to human perceptive organs, and memory to memory. At even lower levels, logic gates are compared to neurons and wires are compared to neuron connections. While such an understanding is sufficient for certain purposes, it may be inadequate for the understanding of natural intelligence emerged from the human brain.

Building a Brain-informatics Portal:
Building a brain-informatics portal is, in fact, to develop a data mining grid centric multi-layer grid system on the wisdom Web for multi-aspect data analysis. The wisdom Web [37,38] and Grid computing [9,14] have provided the ideal infrastructures, platforms, and technologies for building such a brain-informatics portal to support cognitive/brain scientists in multi-aspect analysis in multiple, large-scale data sources. We need to study experimental cognitive neuroscience, data mining, intelligent agents, data and knowledge grids, the semantic Web and wisdom Web in a unified way.
Middleware, as a new platform, is required to cope with multiple huge, distributed data sources for multi-aspect analysis in building a brain-informatics portal on the wisdom Web. It is necessary to create a grid-based, organized society of data mining agents, called a data mining grid on the Grid computing platform (e.g. the Globus toolkit) . This means:
– developing various data mining agents for various services oriented multi aspect data analysis.
– organizing the data mining agents into a Grid with multiple layers such as data-grid, mining-grid, and knowledge-grid, under the OGSA (Open Grid Services Architecture) that firmly aligns with service-oriented architecture and Web services, in order to understand the user's questions, transform them to data mining issues, discover the resources and information about the issues, and get a composite answer or solution.
– using a conceptual model with three levels of workflows, namely data-flow, mining-flow, and knowledge-flow, corresponding to the three-layer Grid, respectively, for managing data mining agents for multi-aspect analysis in distributed, multiple data sources and for organizing the dynamic, status-based processes of brain informatics study.

The data mining grid is made of many smaller components that are called data mining agents. Each agent by itself can only do some simple thing. Yet when joining these agents on the Grid, more complex tasks for brain informatics study can be carried out. Ontologies are also used for the description and integration of multiple human brain data sources and grid-based data mining agents in data mining process planning [30, 90, 92]. It is necessary to provide:

– A formal, explicit specification for the integrated use of multiple human brain data sources in a semantic way.
– A conceptual representation about the types and properties of data/knowledge and data mining agents, as well as relations between data/knowledge and data mining agents.
– A vocabulary of terms and relations to model the domain and to specify how to view the data sources and how to use data mining agents.
– A common understanding of multiple human brain data sources that can be communicated among grid-based data mining agents.

As two related emerging fields of research, Web Intelligence and Brain Informatics mutually support each other. Their synergy will yield profound advances in the analysis and understanding of data, knowledge, intelligence and wisdom, as well as their relationships, organization and creation process. When WI meets BI, it is possible to have a unified and holistic framework for the study of machine intelligence, human intelligence, and social intelligence. BI emphasizes on a systematic approach for investigating human information processing mechanisms, including measuring, collecting, modeling, transforming, managing, and mining multiple human brain data obtained from various cognitive experiments by using fMRI and EEG. In other words, human brain is

regarded as an information processing system, and a systematic study including the investigation of human thinking centric mechanisms, the design of cognitive experiments, data management, and data analysis. Multi-aspect analysis in multiple human brain data sources based on a data brain model is an important methodology in BI. The proposed methodology attempts to change the perspective of cognitive/ brain scientists from a single type of experimental data analysis towards a holistic view at a long-term, global field of vision to understand the principles and mechanisms of HIPS. New generations of WI research and development need to understand multi-nature of intelligence in depth. The recently designed instrumentation (fMRI etc.) and advanced IT are causing an impending revolution in both WI and BI, making it possible for us to understand and develop human-level Web intelligence.

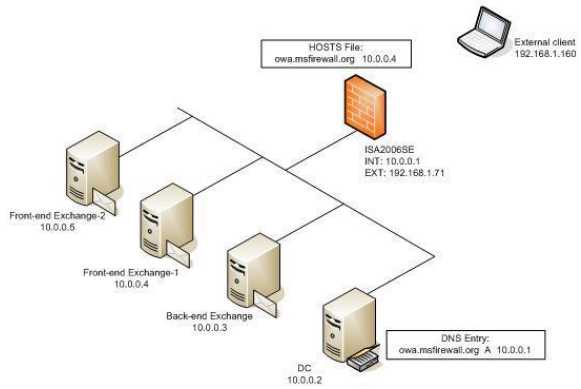
VII. "WEB INTELLIGENCE TECHNOLOGY & ITS APPLICATION":

Web intelligence technology is combination of web analytics which examines how the visitors view and interact with the site's pages and features the development of this technique started from the introduction of new internet based services like online banking, online shopping, e-commerce, etc., the further enhanced development of web intelligence technology by high speed networks, system security, secure transactions will enable safer means of using the features offered by web and also can earn the peoples reputation which contributes to global market.

i. WEB INTELLIGENCE SYSTEMS:

High frequency signal transfer systems, automatic detection systems for finding system malfunction in certain aspects. Even high-end flaw determining or failure of the functioning of certain parts of aircraft or submarine or a ship which sends a warning signal automatically within seconds of malfunction. Various security systems involved in opening and closing of security doors in banks, laser detection system etc. were few of the intelligence systems used in present.

ii. WEB FARMING:

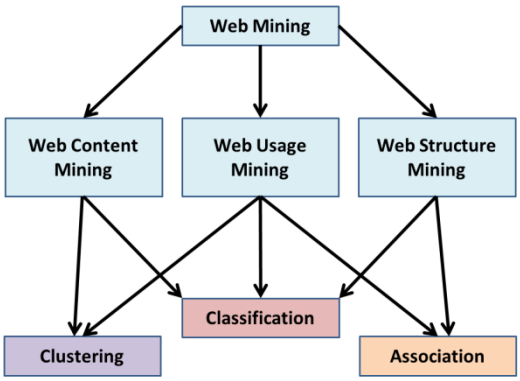


Web farming also known as the server farming or server cluster is the collection of computer servers usually maintained by an enterprise to accomplish server needs far beyond the capability of one machine. Server farms often have backup servers, which can take over the function of the primary servers in the event of the server power failure. Server farms are typically collocated with the network switches and/or routers which enable the communication between the different parts of the cluster and the users of the cluster. The computer routes, power supplies and related electronics which are typically mounted on 19-inch racks in a server room or data centre. Server farms are commonly used for the cluster computing. Many modern super computers comprise giant server farms of high speed processors connected by either gigabit Ethernet or custom inter connects such as infiniband or myrinet server farms are most increasingly being used instead of or in addition to mainframe computers.

iii. WEB SPIDERS:

When most people talk about internet search engines, they really mean World Wide Web search engines. Before the web became the most visible part of the internet, there were already search engines in place to help people find information on the net. Programs like ,goper and ,archie kept indexes of the files stored on servers connected to the internet, and dramatically reduced the amount of time required to find programs and documents. In the late 1980's, getting serious value from the internet meant knowing how to use gopher, archie, venoria and the rest.

iv. WEB MINING:



Web mining is the application of the data mining techniques to discover patterns from the web. According to the analysis targets, web mining can be divided into three categories Web usage mining, Web content mining and Web structure mining, as follows:

- **WEB USAGE MINING:**
  Web usage mining is the process of extracting useful information from server logs i.e., users history. Web usage mining is the process of finding out what users are looking for internet. Some users might be looking at only textual data whereas some others might be interested in multimedia data.

- **WEB STRUCTURE MINING:**
  Web structure mining is the process of graph theory to analyze the node and connection structure of the web site. According to the type of web structural data, web structure mining can be divided into two kinds:

- **WEB CONTENT MINING:**
  Mining, extraction and integration of useful data, information and knowledge from web page contents, the heterogeneity and the lack of structure that permeates much of ever expanding information sources on the world wide web, such as hypertext documents, makes automated discovery, organization, and search and indexing tools of internet and the world wide web such as Lycos, Alta Vista, Web Crawler, ALIWEB, MetaCrawler and others provide some comfort to the users, but they do not provide any structural information nor categorize, filter, or, interpret document.

Thereby the artificial intelligence technology in the field of web has reached greater heights at present and also there should be precautions to illegal usage of precious resources in web. The web intelligence technology is developing day by day and also trying its best to offer quality services to the users all over the world. Even super computers similar to the Human brain can also be developed within few years by 2020 it is expected to be deployed.

VIII. TRENDS AND CHALLENGES OF WI RELATED RESEARCH AND DEVELOPMENT:

Web Intelligence presents excellent opportunities and challenges for the research and development of new generation Web-based information processing technology, as well as for exploiting business intelligence. With the rapid growth of the Web, research and development on WI have received much attention. We expect that more attention will be focused on WI in the coming years. Many specific applications and systems have been proposed and studied. Several dominant trends can be observed and are briefly reviewed. E-commerce is one of the most important applications of WI. The e-commerce activity that involves the end user is undergoing a significant revolution. The ability to track users' browsing behavior down to individual mouse clicks has brought the vendor and end customer closer than ever before. It is now possible for a vendor to personalize his product message for individual customers at a massive scale. This is called targeted marketing or direct marketing Web mining and Web usage analysis play an important role in e-commerce for customer relationship management (CRM) and targeted marketing. Web mining is the use of data mining techniques to automatically discover and extract information from Web documents and services. Zhong et al. proposed a way of mining peculiar data and peculiarity rules that can be used for Web-log mining. They also proposed ways for targeted marketing by mining classification rules and market value functions. A challenge is to explore the connection between Web mining and the related agent paradigm such as Web farming that is the systematic refining of information resources on the Web for business intelligence. Text analysis, retrieval, and Web based digital library is another fruitful research area in WI. Topics in this area include semantics model of the Web, text ming, automatic construction of citation. Abiteboul et al. systematically investigated the data on the Web and the features of semi-structured data. Zhong et al. studied text mining on the Web including automatic construction of ontology, e-mail filtering system, and Web-based ebusiness systems. Web based intelligent agents are aimed at improving a Web site or providing help to a user. Liu et al. worked on e-commerce agents . Liu and Zhong worked on Web agents and KDDA (Knowledge Discovery and Data Mining Agents).

IX. FUTURE OF WEB INTELLIGENCE:

We hope that new granular sets and new granular logical systems with four or more membership functions will be developed in the future to handle Web uncertainty effectively

and fundamentally. Web uncertainty is a long-term challenging problem related to many Web applications like semantic Web, Web mining, Web knowledge discovery, Web agents, Web search engines, Web security, e-Commerce, e-Business, etc. To handle Web uncertainty, we need to develop relevant intelligent Web technology such as CWI and GWI. Importantly, we need to continue to create new granular sets such as neutrosophic sets to try to solve Web uncertainty effectively. Web uncertainty is a difficult long-term problem. So we need to use different intelligent techniques together for this complicated problem. Hybrid Web Intelligence (HWI), a broad hybrid research area, uses Al, CI, BI (Biological Intelligence) and WT to build hybrid intelligent Web systems to handle Web uncertainty effectively and efficiently. In the future, HWI will have a lot of intelligent Web applications under uncertainty. Main HWI applications include (1) intelligent Web agents for e-Applications such as e-Commerce, e-Government, e-Education and e-Health, (2) intelligent Web security systems such as intelligent homeland security systems, (3) intelligent Web bioinformatics systems, (4) intelligent grid computing systems, (5) intelligent wireless mobile agents, (6) intelligent Web expert systems, (7) intelligent Web entertainment systems, (8) intelligent Web services, (9) Web data mining and Web knowledge discovery, (10) intelligent distributed and parallel Web computing systems based on a large number of networked computing resources, ..., and so on.

CONCLUSION:

Perspectives of WI As a new branch of research, Web Intelligence exploit Artificial Intelligence (AI) and Information Technology (IT) on the Web. On the one hand, it may be viewed as applying results from these existing disciplines to a totally new domain. On the other hand, WI may also introduce new problems and challenges to the established disciplines. WI may also be viewed as an enhancement or an extension of AI and IT. It remains to be seen if WI would become a sub-area of AI and IT or a child of a successful marriage of AI and IT. However, no matter what happens, studies on WI can benefit a great deal from the results, experience, success and lessons of AI and IT.

REFRENCES:

[1] REFERENCES Research Challenges and Trends in the New Information Age Y.Y. Yao1, Ning Zhong, Jiming Liu, and Setsuo Ohsuga , IEEE.

[2] Web Intelligence: New Frontiers of Exploration Yiyu (Y.Y.) Yao Department of Computer Science, University of Regina, saskatchewa, IEEE.

[3] www.wikipedia.com/web_intelligence.

[4] Education and the Semantic Web Vladan Devedzic, Department of Information Systems and Technologies, FON – School, of Business Administration, University of Belgrade.

[5] Computational Web Intelligence and Granular, Web Intelligence for Web Uncertainty, Yan-Qing Zhang, Member, IEEE Division of Computer Science.

Author Profile

Rahul Pareek received the B.C.A from Lucky Institute of professional studies, Jodhpur (Raj.) and pursuing M.C.A degrees from Sinhgad Institute of Management, Pune in 2012 and 2015, respectively. He has also done Microsoft Certification in .Net 3.5 and is MCTS.