IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-6, Issue-1 E-ISSN: 2347-2693

Knowledge Analytics in Cloud Centric IoT Vicinities

Nilamadhab Mishra[1*], Kindie Alebachew[2] , Bikash Chandra Patnaik[3]

[1*] Post Graduate Teaching & Research Dept., 09 School of Computing, Debre Berhan University, Debre Berhan 445, Ethiopia
[2] 09 School of Computing, Debre Berhan University, Debre Berhan 445, Ethiopia
[3] Gandhi Institute of Engineering and Technology, Orissa, India

[*]*Corresponding Author: nmmishra77@gmail.com, Tel.: +251-906638096*

Available online at: www.ijcseonline.org

Received: 30/Dec/2017, Revised: 06/Jan/2018, Accepted: 22/Jan/2018, Published: 31/Jan/2018

Abstract— The rapid increasing of real-time applications in today's IoT (internet of things) world progressively lead to several problem issues such as, data volume, velocity, variety, and value. The study reveals that around 80% data in today's IoT world are unstructured and needs an extensive knowledge exploration framework to turn the massively produced data into cognitive values (knowledge) of goldmines. In contemporary IoT vicinities, the time to get knowledge is very slow and the applicability of the knowledge is very poor, so the knowledge researchers start looking new framework that deals with the problems of semantic knowledge analytics and inference. In a typical semantic knowledge analytic scenario, context specification, rule specification, and frame specification may be used to define the structural relationships of knowledge, where the contexts, rules, and frames are stored as specification of framework and sub-framework. In this work, we investigate the viability of a context and rule Analytic framework (CORA-framework) for trustful knowledge analytic and inference in the cloud centric IoT vicinities. We also investigate a knowledge inference case in order to estimate the probabilistic error analysis based on outlier prospect and the knowledge analytic precision based on the root mean square error prospect. The analysis and discussion suggest implementing an outlier analytic mechanism to increase CORA-framework accuracy with estimating the error prediction outcome.

Keywords— IoT (internet of things), context and rule specification, outlier analytic, IoT knowledge analytic, cloud

I. INTRODUCTION

The IoT (internet of things) evolutionary network connects people, processes, places, and things to internet for communication in and around the universe. The IoT objects focus both physical and logical things. The logical things include process, framework, applications, software, and program, and the physical things include people, places, physical entities, and devices. The data of such physical and logical things constitute a comprehensive IoT data base, where the structured, semi-structured, and unstructured data are available [1]. In an IoT data base, ERP and CRM data are considered as structured data, XML data are normally considered as semi-structured data, and email documents, social web contents, pdf, ward, rich text documents are considered as un-structured data. The research finds that, the IoT data base has around 80 % of unstructured data and has no pre-defined data models. Such un-structured data are textual, graphics, video, and symbols oriented. The spatial-temporal databases having the facts or events with time-stamps are also a part of IoT database.

The rapid increasing of IoT applications in today's IoT world progressively lead to several problem issues such as, data volume, velocity, varieties, and value. Analyzing and inferencing cognitive values (knowledge) from large scale IoT data base in a real-time basis is more challenging day by day with the extreme growing of volume, and varieties data that are associated with numerous IoT applications[2],[3]. Such IoT knowledge analytics and inference face a number of real-time problems such as, managing heterogeneous knowledge, transforming varieties data into knowledge, transforming knowledge into actions, transforming actions into cognitive decisions, and tuning the cognitive decisions to coordinate the IoT applications [4],[5].

Now the problem is - where to implement the IoT knowledge analytics and inference mechanism to accomplish the solution of the target problems. In figure 1, we sketch cloud centric IoT vicinities, where the knowledge analytics and inference mechanism can be effectively implemented in knowledge cloud [6]. From IoT vicinities large and bulk amount of data can be transferred to data cloud through the communication gateways and accessed by the data users. Inside the knowledge cloud, a cognitive specification based knowledge analytics and inference framework must be implemented so as to access the right knowledge by right knowledge users at right time to regulate the application intellects.

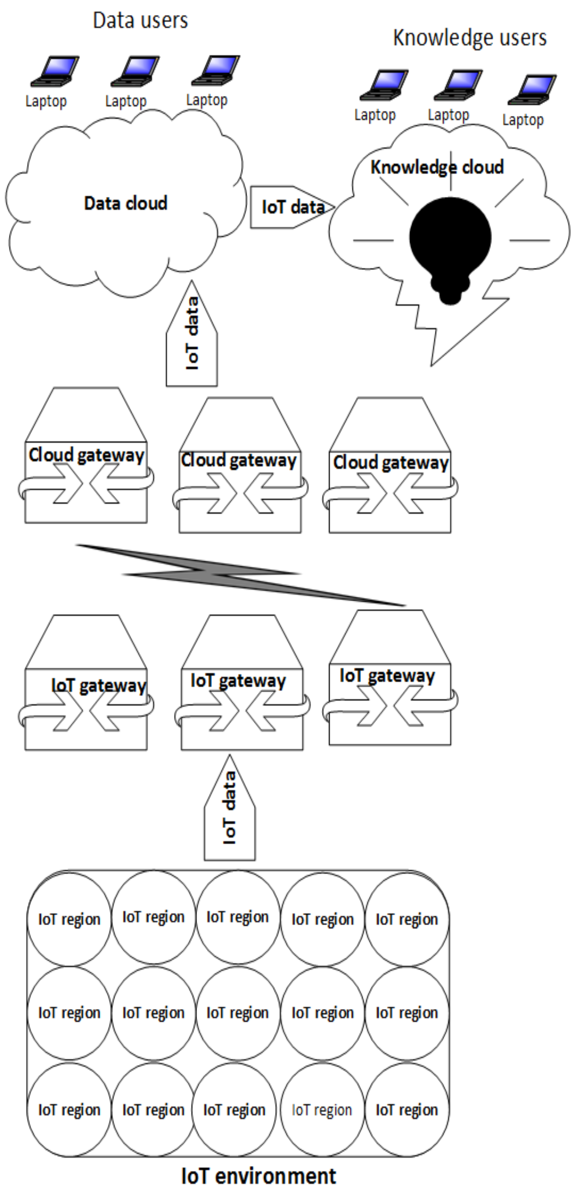

Figure 1. Cloud centric IoT vicinities

The observation has been made that, in a contemporary IoT vicinities, the time to get knowledge is very slow and the quality and applicability of knowledge is very poor [7],[8]. So the knowledge researchers start looking new technologies and tools that deals with the problems of semantic knowledge analytics. In a typical semantic knowledge analytic scenario, context specification, rule specification, and frame specification may be used to define the structural relationships of knowledge, where the contexts, rules, and frames are stored as specification of framework and sub-framework.

The context specification defines the events in a standard conceptualize form and the context aware IoT knowledge analytics needs an exploration framework to configure the semantic knowledge. The rule specification constitutes a consequence of conditions and actions through a set of linguistic variables and values to define the events in a standard rule specification framework. Thus, the hybridization of context and rule exploration framework gives a concise analytic scenarios through customizing the event specification layouts for easy to comprehend and apply so as to overcome the problems of IoT knowledge analytics [9],[10]. So in our work, we investigate the feasibility of a CORA-framework for effective knowledge analytics and inference in cloud centric IoT vicinities. The CORA-framework is a cognitive specification framework that dynamically analyzes the current contexts and rules in order to determine the mismatched rules and mismatched scenarios from possible inferences.

The empirical outlier analytic mechanism is very challenging to implement in CORA-framework to drive the knowledge analytics and inference operations with greater trust and precisions through minimizing the error probability by reducing the outlier prospect. Very few works have discussed the trust analysis of knowledge analytics and inference operations through outlier analytic mechanisms [11],[12],[13].

The rest of this paper is organized as follows. Section II discusses the CORA-framework organization that is associated with context and rule exploration of semantic knowledge analytics in the cloud centric IoT vicinities. Section III highlights the analysis and discussions of the proposed CORA-framework, where outlier analytic mechanism is used to predict the error probability based on the outlier prospect. Finally section IV concludes this paper along with a future work.

II. CORA-FRAMEWORK ORGANISATION

In this section, we discuss CORA-framework organization along with the flow description that implement the IoT knowledge analytic and inference mechanism to obtain the potential knowledge for numerous real-time applications.

A. CORA Flow description:-

Here we will discuss a flow description to emphasize both rules and target goal as important factors. The purpose of design the work flow (see figure 2) is to focus on analytics and inferences from IoT context-specification that consists of historical events or observations.

The IoT context-specification enables the knowledge mining researchers to extract the possible antecedents that persists linguistic variables and values. Then the fuzzification process is initiated through assigning the approximate fuzzy membership grades to quantify the linguistic values [7]. A rule-specification is used that contains pre-defined rules for the applications. If rules are in-sufficient in accordance with the extracted antecedents, then new rules are added to rule-specification so as to extract the possible consequents. Now, if the possible consequents cannot meet the target goal of application, then the rule-specification base needs to be re-engineered. Finally, the outputs are normalized to obtain the desired knowledge inferences for the application. In the work flow structure, the context and rule-specification are configured based on the application target so as to accomplish the desired goal effectively [14],[15],[16].

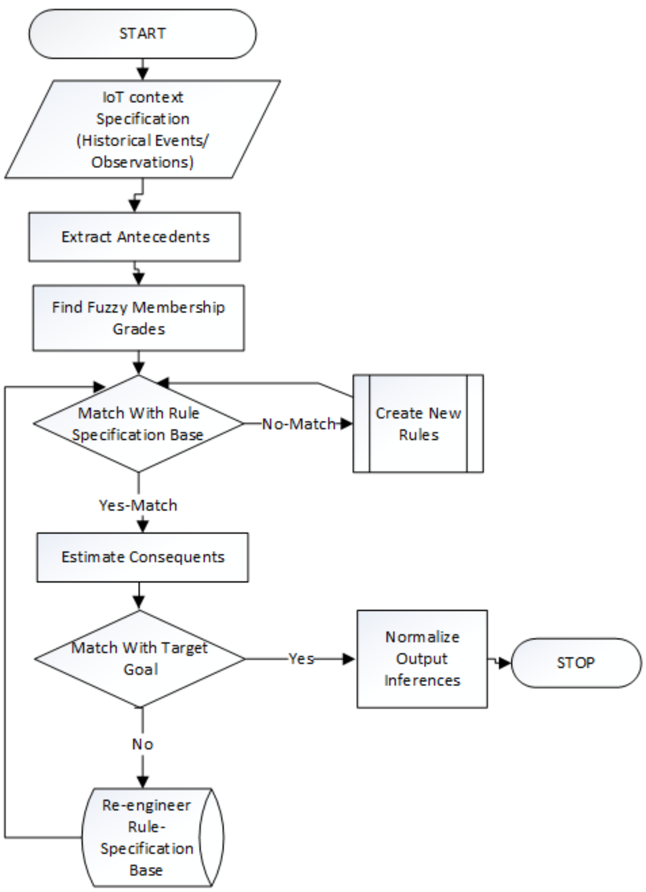

Figure-2. Work flow for analytics and inferences from IoT context-specification that consist of historical events or observations

B. Framework organization:-

In CORA-framework organization, the initial activity is the potential identification and classification of inputs based on the analysis of IoT context-specification of the target application. The layer-1 has a two input classes, in which class-A and class-B inputs consist of valid linguistic variables and values respectively.

In figure-3, we describe a layering architecture of CORA-framework with input class having three linguistic variables and three linguistic values. In layer-2, we define the valid combinations of linguistic variables and values to generate possible antecedents for the application. The current IoT events X at time stamp t "X (t)" fetch to the layer-2 for the weights processing to the next layer. In-between layer-2 and layer-3, the connection weights are assigned stochastically based on the actual analysis of X (t). The layer-3 is a filtration layer that filters the antecedent with the highest stochastic value. The layer-4 is a rule-layer, where the rules are derived from the rule-specification base of the target application. Each antecedent must interact with individual rule for potential implications.

The layer-5 preserves the valid consequents based on the rules. The layer-6 is a normalization and de-fuzzification layer, where the consequents are normalized to obtain the computed inference outputs at time stamp t, i.e. Y (t).

In a supervised system, the computed outputs Y (t) are matched with the target outputs T at time stamp t T (t), in order to estimate the probable system errors that should be acceptably low to main better system accuracy. To achieve better system accuracy, an effective mechanism should be used that adjusts the connection weights in between the layers so as to lead an accurate result.

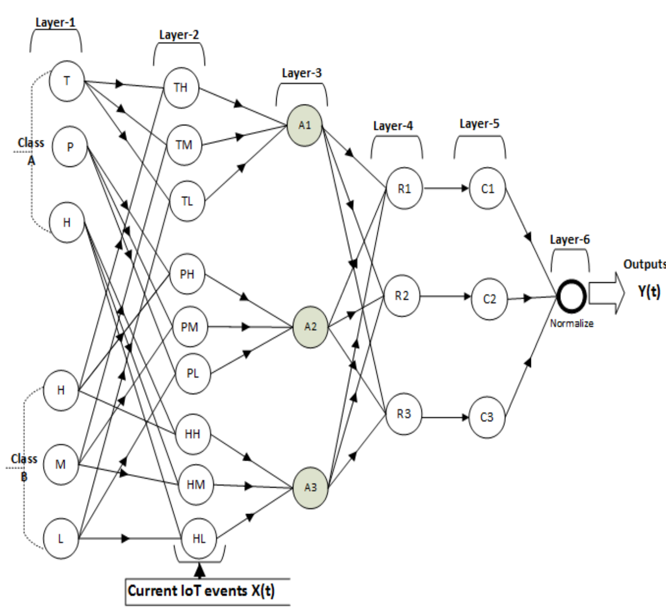

Figure-3. Layering of CORA-framework with input class having three linguistic variables and three linguistic values

III. ANALYSIS AND DISCUSSION

In this section, we discuss the contexts and rules analytics mechanism for CORA-framework and some real-time application data are trace to the framework to perform the knowledge analytic and inference operation so as to predict the output errors.

A. Knowledge inference case analytics:-

In this case analytics, we explore huge debt analysis context cases for the framework. The case analyses several customer's debt ratio (DR) based on the revenue utilizations (RU). In this analysis, a huge number of customer's instances are generated as input data instances to the framework. The input customer's instances are sketched in figure-4 as described. We find following statistical analysis over 90645 customer's instances. In revenue utilization analysis, min =0, max= 1, average= 0.313, and standard deviation= 0.336. In debt ration analysis, min= 0, max= 1, average= 0.308, and standard deviation= 0.223.

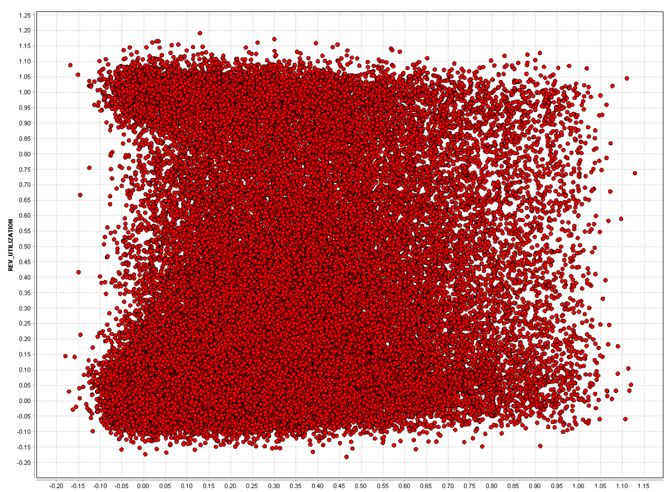

Figure-4. Input analysis of large scale customer's instances (X: Debt_ratio; Y: axis- Rev_utilization)

The analysis finds three analytic rules for the framework.
Rule 1. If RU = "high" then DR = "high"
Rule-2. If RU = "low" then DR = "low"
Rule 3. If RU = "average" then DR= "average"

For three analytic rules, the fuzzy membership grades are defined as follows.

i. μ (low) (x): x ϵ X and 0.00⩽ μ (low) (x) ⩾ 0.40]; iff x is a customer's instance and X is a customer data set.

ii. μ (Avg) (x): x ϵ X and 0.30⩽ μ (Avg) (x) ⩾ 0.70]; iff x is a customer's instance and X is a customer data set.

iii. μ (high) (x): x ϵ X and 0.60⩽ μ (high) (x) ⩾ 0.10]; iff x is a customer's instance and X is a customer data set.

The input customer's instances are traced to the CORA-framework to predict the probabilistic errors based on the quantity of outliers present in the instances. In table-1, we estimate the outlier prospect to mean square (MS) error probability from the input customer instances so as to predict the overall system accuracy. The figure-5 analyses the error probability based on the outlier prospect of the CORA-framework. The analysis states that the mean square error probability (MSEP) increases with increasing the outlier prospect. The MS error probability = 1 when outlier prospect > 0.80. So, in order to increase the accuracy of CORA-framework, the outlier prospect should acceptably low for efficient and trustful knowledge analytics and inferences in the knowledge cloud of cloud centric IoT vicinities.

Table-1. Probabilistic error analysis report

Outlier prospect	MSEP
0.01	0.2788
0.03	0.4545
0.05	0.6359
0.10	0.8034
0.20	0.9139
0.30	0.9546
0.40	0.9776
0.50	0.9902
0.60	0.9961
0.70	0.9989
0.80	0.9998

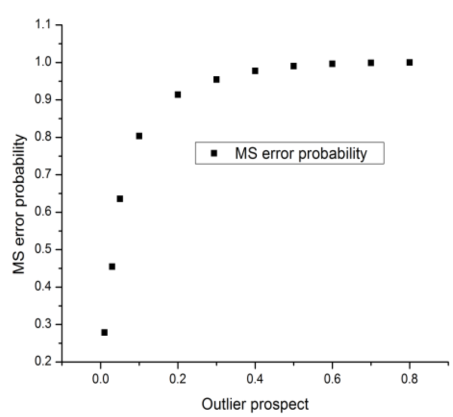

Figure-5. Error probability analysis based on outlier prospect

In Figure 6, we present a KAP-error analysis result in terms of estimating the errors (RSME) for the CORA-framework. Our analysis states that, in a rule based context scenario KAP-error is directly propositional to RSME value.

$RSME_{value}$ α KAP-error --------------- (1)

The equation-1 indicates that, higher the RSME values higher are the KAP-errors and vice-versa.

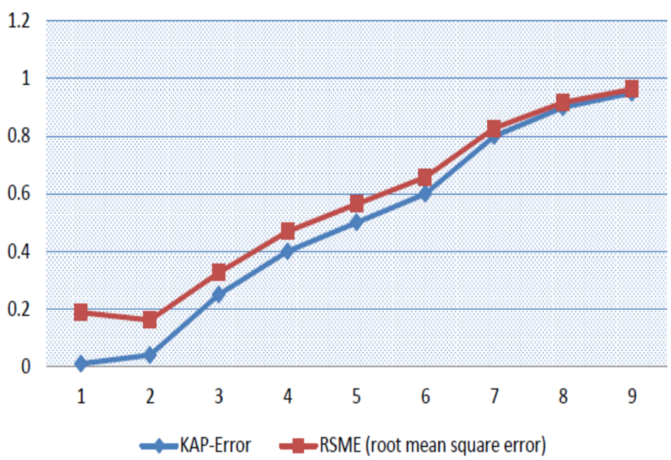

Figure-6. Knowledge Analytic Precision error based on RSME

The convergences of statistical and computational learning mechanisms have been researched to deal with the knowledge analytic problems. Knowledge analytic implements are also used for analysing and exploring various operational tasks associated with the big data submissions, such as-data transformation and analysis, data mining, knowledge discovery, semantic knowledge explorations, structural analysis, and many more. The machine learning technics are implemented in many areas of knowledge discovery and semantic knowledge analytics to explore the application intelligence.

In this work, we analysed the precise characterization of the knowledge analytic problem through a case scenario, context in which the analytic rule is implemented, context from which the specified inferences are derived, and the feasibility of applying the analytics to cloud centric IoT vicinities.

IV. Conclusion and Future Scope

In this work, we discuss cloud centric IoT vicinities, in which knowledge analytic and inference mechanism is implemented in the knowledge cloud. To enable the knowledge analytic and inference from IoT data base, we propose a CORA-framework that includes six layers to perform different knowledge analytic operations. We highlight a work flow structure for analytics and inferences from IoT context-specification that consists of historical IoT events or observations. The both rule-specification and context-specification are emphasized while analysing the outlier analytic mechanism. We perform a KAP-error analysis through a business case analytic for our proposed CORA framework to empirically estimate the MSE and RMSE. The most important factor is to configure the rule-specification and context-specification based on the target application so as to accomplish the desire goal. In the analysis and discussion section, we discuss an outlier analytic mechanism for the said framework and some real-time application data are traced so as to predict the MS error probability with respect to outlier prospect with an aim for trustful knowledge analytics and inferences in the knowledge cloud of cloud centric IoT vicinities.

In future issue, we will implement new learning mechanism to design new application specific algorithms for the knowledge analytic and inference set-ups for cloud centric IoT vicinities.

Acknowledgment

The authors would like to express thanks to the Post Graduate Teaching & Research Dept., at School of Computing, Debre Berhan University, Ethiopia for supporting this research.

References

[1] Mishra, Nilamadhab, Chung-Chih Lin, and Hsien-Tsung Chang. "A Cognitive Oriented Framework for IoT Big-data Management Prospective."Communication Problem-Solving (ICCP), 2014 IEEE International Conference on. IEEE, 2014.

[2] S.J. Nasti, M. Asgar, M.A. Butt , "Analysis of Customer Behaviour using Modern Data Mining Techniques", International Journal of Computer Sciences and Engineering, Vol.5, Issue.12, pp.64-66, 2017.

[3] Zarko, Ivana Podnar, et al. "IoT data management methods and optimisation algorithms for mobile publish/subscribe services in cloud Environment. "Networks and Communications (EuCNC), 2014 European Conference on. IEEE, 2014.

[4] Gubbi, Jayavardhana, et al. "Internet of Things (IoT): A vision, architectural elements, and future directions." Future Generation Computer Systems 29.7 (2013): 1645-1660.

[5] Anantharam, Pramod, Payam Barnaghi, and Amit Sheth. "Data Processing and Semantics for Advanced Internet of Things (IoT) Applications: modeling, annotation, integration, and perception." Proceedings of the 3rd International Conference on Web Intelligence, Mining and Semantics. ACM, 2013.

[6] Mishra, Nilamadhab, Hsien-Tsung Chang, and Chung-Chih Lin. "An IoT Knowledge Re-engineering Framework for Semantic Knowledge Analytics for BI-services." Mathematical problems in engineering, vol. 2015, Article id-759425, 12 pages (2015).

[7] Dagnino, Aldo, and David Cox. "Industrial Analytics to Discover Knowledge from Instrumented Networked Machines." Proceedings of the 26th International Conference on Software Engineering and Knowledge Engineering (SEKE'14), Vancouver, Canada. 2014.

[8] Chang, H. T., Mishra, N., & Lin, C. C., IoT big-data centred knowledge granule analytic and cluster framework for BI applications: a case base analysis. PloS one, 10(11), 2015.

[9] Mishra, N., Chang, H. T., & Lin, C. C. Sensor data distribution and knowledge inference framework for a cognitive-based distributed storage sink environment. International Journal of Sensor Networks, 26(1), 26-42,2018.

[10] Mishra N,. "In-network Distributed Analytics on Data-centric IoT Network for BI-service Applications", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN: 2456-3307, Volume 2, Issue 5, pp.547-552, September-October.2017.

[11] Patnaik, B. C., & Mishra, N. "A Review on Enhancing the Journaling File System", Imperial Journal of Interdisciplinary Research, 2, no. 11 (2016)

[12] Chang, H. T., Yu-Wen Li., & Mishra, N. "mCAF: A Multi-dimensional Clustering Algorithm for Friends of Social Network Services", Springer Plus, 2016.

[13] Chang, H. T., Liu, S. W., & Mishra, N. "A tracking and summarization system for online Chinese news topics", Aslib Journal of Information Management, 67(6), 687-699,2015.

[14] Mishra, N., Lin, C. C., & Chang, H. T. "A Cognitive Adopted Framework for IoT Big-Data Management and Knowledge Discovery Prospective", International Journal of Distributed Sensor Networks, 2015.

[15] Mishra, N., Lin, C. C., & Chang, H. T. "Cognitive inference device for activity supervision in the elderly", The Scientific World Journal, 2014.

[16] Mishra, N., Chang, H. T., & Lin, C. C. "Data-centric Knowledge Discovery Strategy for a Safety-critical Sensor Application", International Journal of Antennas and Propagation, Article ID 172186, 11 pages, 2014. doi:10.1155/2014/172186.

Authors Profile

Dr. Nilamadhab Mishra is currently an Assistant Professor in Post Graduate Teaching & Research Dept., at School of Computing, Debre Berhan University, Ethiopia. He accomplishes his PhD in Computer Science and Information Engineering from Chang Gung University, Taiwan. He moreover publishes numerous peer reviewed researches in Thomson Reuter's ranked SCI journals & IEEE conference proceedings, and serves as reviewer and editorial member in peer reviewed Journals and Conferences. Dr. Mishra's research areas focus on Network Centric Data Management and Knowledge Discovery, IoT Data Science and Knowledge Analytics, Business Intelligence, and Cognitive Applications exploration. He has 15 years of Academic Teaching and Research Experience.

Mr. Kindie Alebachew is currently working as a lecturer in College of computing at Debre Berhan University, Ethiopia. He completed his masters of Science in Information Science at Addis Ababa University, Ethiopia. His Research interests comprise Machine Learning, AI, and Data Science. He has 5 years of teaching and research Experience.

Mr. Bikash Chandra Patnaik is currently working as a professor in Gandhi Institute of Engineering and Technology, Orissa, India. He completed his masters of technology in Computer Science at Utkal University, India. His Research interests comprise Machine Learning and Data Science. He has 20 years of academic teaching and research Experience.

