IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 3, September 2024, pp. 2546~2556
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i3.pp2546-2556

Framework for contextual consulting practices in adherence for decentralized data-driven decision making

Vijay Kumar Pandey, Neeraj Kumar Rathore, Narayan P. Bhosale
Department of Computer Science, Faculty of Computronics, Indira Gandhi National Tribal University, Amarkantak, India

Article Info

Article history:

Received Sep 2, 2023
Revised Nov 9, 2023
Accepted Jan 6, 2024

Keywords:

Consulting practice
Data-driven
Decentralization
Industry 4.0
Knowledge discovery

ABSTRACT

With the rising adoption of technological advancement and industry-based automation standards, the area of consulting firms is gradually evolving to keep up this pace towards incorporating sophisticated analytical operation for facilitating value-added consulting services. Review of existing practices of consulting firm shows increasing adoption of analytical process which leads to complex form of operation towards knowledge discovery of consulting data. Hence, this manuscript introduces a framework of contextual consulting practices where the core idea is to incorporate a baseline structure of knowledge discovery associated with consulting data in adherence of industry 4.0 automation standards. The framework takes the input of streamed consulting diversified data governed by a template-based entry-points where the consulting data is subjected to series of transformation operation that not only preprocess the consulting data but also optimizes the data to enhance its data quality. The study model is implemented in MATLAB considering an extensive analytical framework towards data-driven decision making and decentralization to exhibit proposed model to offer better analytical performance in contrast to existing study models.

Corresponding Author:

Vijay Kumar Pandey
Department of Computer Science, Faculty of Computronics, Indira Gandhi National Tribal University
Amarkantak, Madhya Pradesh 484887, India
Email: vijay.reserach4@gmail.com

1. INTRODUCTION

With the evolving trends of contextual consulting practices, there is a rising scope as well as competition towards offering the best form of services irrespective of any domain of consulting industry [1]. Basically, contextual consulting directs to facilitating of highly customized recommendation based on catering up specific challenge or demand of the customer (or enterprises) [2]. Different from generalized consulting practices, there is a gaining popularity and demand of contextual consulting as it focuses uniquely on ever clients considering their constraints, goals, culture, and industry [3]. An effective consulting practices demands manifold attributes e.g., ethical consideration, result measurement, continuous learning, communication skills, flexibility, collaboration, log-term perspective, customized solution, and comprehensive analysis [4]–[7]. It is important to understand that there is an involvement of multiple number of consultants, both in form of individual and a firm, working in collaborative manner in order to cater up the demands of its client [8]. With the advent of technological advancement, contextual consulting practices is caried out using manifold technologies in order to aggregate, analyze, as well as analyze data considering a particular organization or business context. Out of all the current form of technologies, business intelligence tools as well as data analytics are reported to be frequently adopted owing to its simplified and yet powerful analytical capabilities [9]. Various analytical approaches e.g., machine learning, clustering, time-series analysis, and regression analysis is deployed in order to extract deeper knowledge associated with complex form of consulting data [10]. As contextual consulting usually consists of various qualitative and quantitative data, comments from internal communication, social media, and client reviews. Therefore, natural language processing (NLP) is adopted by the consultants to simplify the task of understanding the data [11]. Apart from this, adoption of artificial intelligence (AI) and deep learning is also gaining pace in advanced data analytics which can offer significant performance improvement in various internal processing of contextual consulting practices [12]–[15]. Along with AI usage, there are also reported cases where corporates use various cloud platforms e.g., Google Cloud, Azure, and Amazon web services in order to harness the analytical capabilities, processing power, and storage facilities [16]. Existing consulting corporates also uses various collaborative tools and project management software in order to incorporate highly synced coordination among the client and consulting team [17]. Apart from technological usage, existing consulting practices are also reported to use geospatial analytics [18], social listening tools [19], blockchain technology [20], augmented/virtual reality [21], simulation software [22], and dashboard and reporting tool [23]. A closer look into the evolution and adoption practices of the above-mentioned tools and technologies shows that contextual consulting practices has already moved to a higher level of verticals of technological advancement. It can also be noted that adoption of such tools and technologies would soon automate the contextual consulting practices to next level in adherence of industry 4.0 standard that is characterized by internet-of-things, automation, data analytics, and digital technologies [24]. The contribution of industry 4.0 is towards improving the capability of a consultant to facilitate highly customized solution to their clients along with potential scope of understanding the business dynamics [25]. Adoption of industry 4.0 standard in contextual consulting practice offers consultants with following benefits viz: data-driven insights, real-time monitoring, predictive analytics, customization and personalization, process optimization, supply chain management, collaborative platforms, innovation, and disruption, change management, global connectivity, and risk management [26], [27].

There are various related works where industry 4.0 standards have been attempted to be incorporated in the form of varied analytical operation in consulting practices. Adoption of industry 4.0 standards and deep learning is noted for study massive size of industry data [28]. Sato *et al.* [29] presented a digital triplet concept towards improving the production system and decision making by leveraging the consulting process model. Adoption of machine learning is reported where the agenda is towards improving the organizational performance using regression tree method [30]. The idea is to improve collaboration and outsourcing within a firm. Deep learning approach is adopted to evaluate the emotional intelligence of an employee [31]. They have used self-organizing map (SOM) and convolution neural network (CNN). Nurek and Michalski [32] have integrated social network with machine learning in order to investigate an internal structure of an organization. Varied machine learning models e.g., support vector machine (SVM), neural network (NN), and decision tree (DT), have been used as collective classification algorithm. Further adoption of machine learning is reported with a target to optimize the process of knowledge management within a business model [33]. Ansari [34] have developed a model using deep unsupervised learning (USL) method of SOM in order to investigate market segmentation. An effective discussion is carried out towards harnessing capabilities of big data approach for investigating growth in learning cultures within an organization [35]. Adoption of contextual embedding is reported where long short-term memory (LSTM) have been used for analyzing textual data [36]. Apart from this, it is also noted that organizational data has been investigated from text-mining approaches too as noted in [37]–[41].

However, there are various challenges associated with the implementation of principles of industry 4.0 in contextual consulting practices inspite of its beneficial features. Some of the potential issues that poses as a research problems are: i) there is a significant issues and complexities associated with technological integration from the wider range of technologies in industry 4.0 e.g., automation system, data analytics, and smart devices; ii) basically, industry 4.0 is an evolving technology in current time, hence effective frameworks and protocol standardization still lacks in existing times; iii) the initial cost of implementing industry 4.0 standard is quite expensive owing to the usage of infrastructure, software, and new hardware; iv) there is a less opportunity for industry 4.0 to seamlessly translate across different organizational cultures and diversities; v) adoption of industry 4.0 leads to massive amount of data generation from multiple sources that can easily lead towards information overload; vi) with the growing speed of technological advancement associated with industry 4.0, there is a risk that some of the currently implemented technologies in important operational business module may become obsolete; vii) the consulting data when managed by existing form of industry 4.0 standard will eventually undergo various set of complex and sophisticated internal analytical operation which may not be generalized for every business domain or every client leading to higher expenditure in project management; viii) adoption of AI may significantly contribute towards potential analytics, however existing system doesn’t report of any simplified or cost effective business intelligence tool dedicated for contextual consulting firms; ix) majority of existing approaches are quite computationally complex in nature although they reports of higher predictive score of accuracy; and x) none of the existing system reports of considering diversified data source of consulting firm nor they have considered cost effective distributed framework for practicising contextual consulting.

Therefore, the prime contribution of the proposed scheme is to introduce a novel framework of contextual consulting practice (FCCP). The methodology adopted is a hybrizationscheme of a novel text mining and simplified knowledge extraction methods are utilized in adherence of industry 4.0 standard. Unlike existing study models, the proposed study accomplishes a good balance between computational efficiency and predictive accuracy. The next section discusses about adopted research method of proposed model.

2. METHOD

The prime goal of the proposed study is to design a novel form of computational framework towards facilitating contextual consulting service adhering to the norms of industry 4.0. Out of manifold characteristic, the proposed scheme emphasizes mainly towards data-driven decision making and decentralization towards the practices of consulting firm. A simplified analytical methodology is adopted for this purpose as shown in Figure 1.

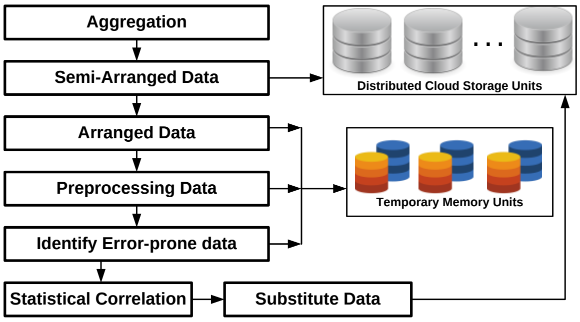

Figure 1. Method for contextual consulting framework

According to Figure 1, there are various sources of consulting data in different geographical regions that are considered to use a unified interface towards data management using a specific form of template. The initial phase of implementation involves preprocessing the aggregated consulting data while a unique approach is being introduced by extracting the part of this aggregated information called as an identifier to be stored in distributed while the part of this aggregated information is stored in temporary buffer. In order to facilitate better form of data driven methodology, the scheme emphasizes on quality of data and undertakes series of operation towards it. The study model generates a semi-arranged data (s_{ad}) and transform it to arranged data (a_d) without the consulting data being stored in permanent storage units over cloud. Further, an algorithm is implemented towards yielding preprocessed data which is a direct form of contextual data. The term contextual consulting data will refer to acquired core analytical meaning of massive raw consulting data that can assists in decision making in reduced time with satisfactory accuracy. Further, the proposed scheme initiates its probe towards identifying the presence of error-prone data. Once such form of artifacts is identified, the proposed scheme contructs a statistical correlation-based matching mechanism between the error data and precise data. As all the information is maintained in the form of rows and columns, the approaches choose to substitute the value corresponding to the row with higher correlation with the cell bearing error-prone data. The final outcome is updated to the cloud storage unit in distributed form that can be access by the stakeholders of consulting part anywhere and anytime. The core novelty of this approach is its simplicity in implementation strategy without adopting complex analytical operation. The overall implementation of the above-mentioned methods is further illustrated in form of algorithm design.

2.1. Algorithm for preprocessing consulting data

This algorithm is responsible for performing preprocessing of the distributed contextual data. Unlike conventional preprocessing operation, the core notion of this algorithm is to offer proper facilitation towards automating incoming stream of data. The study considers acquisition of such forms of large sized data originates from various consulting firms via cloud services that could result in generation of highly unorganized arrangement of consulting data. The study assumes that a specific form of template is maintained from the client-side for user-friendly entry of data while the accumulated data on the server side spontaneously keeps on increasing and is required to be maintained in distributed form for easy accessibility. Hence, the prime idea of this algorithm is to address this issue of sub-optimal arrangement of data in distributed cloud. The steps of this algorithm are shown in Algorithm 1.

Algorithm 1. Preprocessing consulting data
Input: n, α, s, v
Output: p_d
Start
1. For i=1:n
2. $c_d=[\alpha, s, v]$
3. $s_{ad} \rightarrow f_1(c_d)$
4. $a_d \rightarrow f_2(s_{ad})$
5. $p_d \rightarrow f_3(a_d)$
6. End
End

The discussion of steps Algorithm 1 is as: the algorithm takes the input of number of data source (n), identifier (α), separator (s), and value (v) that after processing yields an outcome of preprocessed data (pd). The study considers at data is generated from multiple consulting firms in different geographical location. All the consultanting firms are assumed to adopt a template-based design structure in order to feed the generated data. The proposed study considers a simplified form of template which consists of identifier, separator, and value as shown in Figure 2.

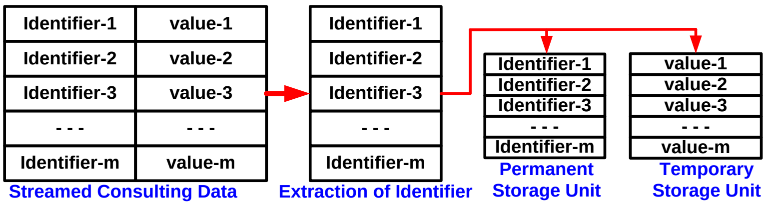

Figure 2. Attributes of template-based source

From Figure 2, it can be seen that there are possibilities of inclusion of various types of identifiers which basically represents the categories of consulting data while individual values represent unique value of each identifier. The corpus also has separator which can basically distinguish identifier from its values and this structure of template is highly useful for identifying and discretizing the respective values. Hence, a matrix of consulting data c_d is formulated (line-2) using these attributes considering all the incoming consulting data n (line-1). The next line of algorithmic operation calls for further optimizing the buffer usage in server by retaining only the identifiers in permanent address while its respective values are still subjected for further analytical processing. For this purpose, an explicit function $f_1(x)$ is formulated (line-3) taking the input argument of matrix from consulting data c_d in order to generate s_{ad}. The function $f_1(x)$ performs two essential tasks:
- It extracts all the identifiers from consulting data matrix c_d and stores in buffer of cloud server following by allocating a unique index for each identifier. However, it doesn't extract any respective values from the c_d matrix.
- After the prior operation is accomplished, the function further assigns a unique index. It is to be noted that these indexes are given separately for assigned identifier and separately for its values. Although, the index maintained in cloud buffer is the final one; however, a second layer of index in assigned for next coming data stream to confirm if the identifiers are of same of different types.

The operations leads to generation of s_{ad} with a confirmation of identifiers and accounted respected values of each identifier. The next part of algorithmic operation is linked with further increasing the accountability of the data by introducing another explicit function f2(x) as exhibited in line-4. This function performs two sequential set of operation as:
- This function takes the input argument of s_{ad} with a confirmed linking of index for identifiers and its respective values and store it in a temporary buffer known as a_d (line-4).

– After the prior task is accomplished, the function uses a grammar tagging where each field within an identifier is cross-matched with the respective grammer in order to derive a contextual meaning of the elements in consulting data.

It is to be noted that proposed scheme doesn't uses any form of lexical dataset (WordNet and SentiNet) in order to perform this task which is much frequent in existing approaches for text mining-based application. The reson behind this is lack of customization and dependencies towards third parties to process the corpus for acquiring contextual information. Hence, the study model offers a novelty by constructing a function which has a customized definition of linguistic grammar and that can frequently be updated by any user without any prior knowledge of operating it thereby making it more user-friendly. Further, the algorithm applies the third explicit function f_3(x) on the recently acquired a_d (line-5) in order to generate a preprocessed data. The operation carried out by function f_3(x) is mainly towards exploring the presence of most dominant and frequently used contextual information present in a_d to confirm it as preprocessed data that is then saved back in the buffer of distributed server permanently with a newly assigned index. Therefore, the distributed server now consists of indexed identifiers and indexed values of contextual data whereas the local server consist already consists of raw consulting data. The overall process is pictorially exhibited in Figure 3.

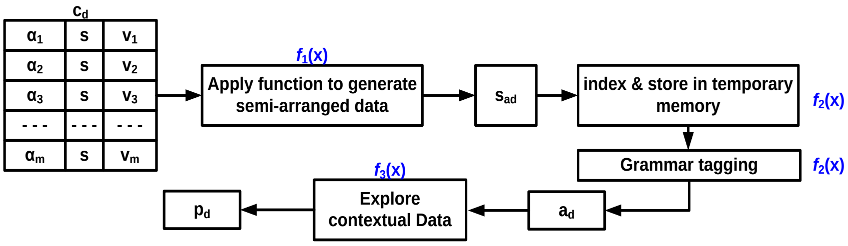

Figure 3. Mechanism for preprocessing data

2.2. Algorithm for optimizing consulting data quality

The prior algorithm is responsible for generating a preprocessed data considering the possibility of unorganized data arrangement owing to massive streaming of online consulting data. However, the algorithm misses the possibility of presence of error-prone data. There is multiple rationale behind the consideration of error-prone data to be associated with preprocessed data obtained from prior algorithm as:

– According to prior algorithm, the original raw consulting data resides in local server while the preprocessed data resides in distributed cloud server. As the complete processing is carried out over a temporary buffer and then the result of preprocessed data is stored back in distributed cloud server, there is a possibility of network-based issues resulting in error-prone data inspite of correct indexing.
– There is also a possibility that user could add a new identifier or edit an existing identifier or delete the older ones. In such case, the updating of the same in distributed cloud server could effectively do the task; however, there is still possibilities of certain values to be overlapped or undertake a form which is beyond recognizable form.

Hence, a new operation is carried out which addresses this problem by formulating a new algorithm. Therefore, this algorithm is responsible for identifying the error-prone data in the contextual preprocessed data followed by replacing it with fair quality data. This algorithm will also form a baseline of automated management of quality data from the massive stream of raw consulting data. The steps of this algorithm are shown in Algorithm 2.

Algorithm 2. Optimizing consulting data quality
Input: p_d
Output: s_d
Start
1. For i=1:n
2. For j=1:p_d
3. $e_d \rightarrow f_4$(j)
4. $s_d \rightarrow f_5(e_d$, j)
5. End
6. End
End

Algorithm 2 takes the input of preprocessed data that after processing yields to substituted data (s_d). The algorithm initiates its operation by considering all the incoming streams of data n (line-1) followed by considering all the recently computed preprocessed data (line-2). Further, the prime contribution of this algorithm is to construct an explicit function $f_4(x)$ that carries out following steps of operation:

- The first task of the function $f_4(x)$ is to evaluate the consistency of the respective values associated with the identifiers by performing a row-wise search operation on the matrix of preprocessed data. The search stops for the cell which has been identified as a pattern of value different from the respective datatype of its parent identifier.
- In order to confirm the presence of error or anomaly value, the proposed scheme initially declares the possible datatypes for each identifier over all the cells retaining respective values. This information serves as a meta-data that is compared with cell consisting of respective values. Upon finding the error-prone data, the function assigns a temporary flag for it to be addressed while the search is continued for the entire corpus.
- The function $f_4(x)$ computes the statistical correlation for all the rows bearing precise as well as error prone data and generates two set of information i.e., i) correlated values of all rows respective to identifier where the error prone data resides and ii) generates a highlight of error prone data e_d (line-3)

In the next line of operation (line-5), another function $f_5(x)$ is constructed which carries out further set of task: i) the function $f_5(x)$ considers the row with error prone data (e_d) and considers its correlated values without considering the error prone data and store it in matrix of correlated data of error prone cell (c_{ed}); ii) the function $f_5(x)$ acquires the coorelated values of other rows of preprocessed data where there is no error (c_{cd}); and iii) finally, the algorithm finds out the higher matching statistically correlation values between error prone data and preprocessed data (line-4) followed by substituting the respective cell value with higher correlation score. Taking an example of Figure 4 to clarify the operation of this algorithm, it shows that cell (r_4, c_3) consists of error-prone data. The algorithm computes correlation for this cell and store its value in c_{ed} matrix, while the algorithm computes correlation of other rows and store them as c_{cd1}, c_{cd2}, c_{cd3}, and c_{cd4}.

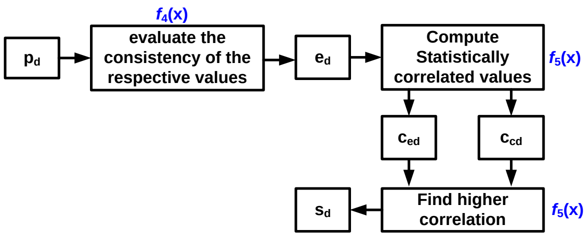

Figure 4. Mechanism of optimizing data quality

All the acquitted statistical correlation is now compared with corelated value of cell with e_d i.e., c_{ed}. For the sake of understanding, sample values are shown in Figure 4 exhibiting higher score for fouth row. Hence, the respective element within cell (r_4, c_3) is now substituted to cell with e_d i.e., cell (r_2, c_3) thereby completing the operation of this algorithm. The next section presents discussion of discussed research method.

3. RESULTS

The proposed FCCP demands to be assessed with massive set of data acquired from consulting firm. For this purpose, the analysis is carried out from Kaggle dataset of consulting survey which consists of 19 fields with both identifier and values [42]. However, the dataset posses less quantity of information and hence it is manually increased to obtain 10,000 fields of information in order to testify the capability of FCCP. Scripted in MATLAB, the proposed model developed 5 sources of origination of consulting data that are streamed over a cloud network in unified interface followed by subjecting it to proposed algorithm. Each originating sources of consulting data is uniformly retained to posses 5000 fields of information in textual format that consists of reviews shared by all involved consultants. The complete model is assessed with respect to four performance parameters viz; accuracy, preprocessing time, data quality, and memoy utilization whose accomplished numerical results were shown in Table 1.

Table 1 showcase that proposed scheme FCCP has been compared with conventional methods of AI as reported in existing studies viz. Classification based approaches using supervised learning (SL) and clustering-based approach using USL. The SL methods consists of SVM while USL method consists of K-means clustering. Although, assessment has been carried out for other segment of SL and USL methods, but better results have been observed for SVM and K-means clustering. The consecutive section further discusses the learning outcomes obtained from the numerical evaluation exhibited in Table 1.

Table 1. Numerical outcome of study

Approaches	Accuracy (%)	Preprocessing time (s)	Data quality (%)	Memory utilization (%)
FCCP	95.6	0.619	98.3	25.7
Classification (SL)	92.7	1.753	96.1	75.6
Clustering (USL)	89.5	3.882	93.7	77.9

3.1. Discussion towards accuracy performance

The parameter of accuracy is used as a numerical value obtained by evaluating the capability of considered knowledge extraction process for acquiring the knowledge from given text ensuring its correctness from ground truth data. Basically, this parameter is essential for testifying the correctness in substitution mechanism introduced by second algorithm of FCCP. The graphical outcome is presented in Figure 5. The outcome showcases that proposed FCCP offers approximately 4.5% of increased accuracy in contrast to existing system of SL and USL. A closer look into Figure 5 showcases that USL offers lesser accuracy which is because of the fact that although USL offers simplified implementation scheme ontext data, yet choosing the right number of clusters in prelimanry level is quite challenging one especially when streams of incoming data is considered. On the other hand, SL methods offers its benefits towards processing high-dimensional data by exploring the optimal hyperplanes; however, it is found to offer higher sensitivity to regularization and kernel parameters which is not only time consuming but also reduces the accuracy. FCCP offers the advantage of statistical correlation-based analysis with dual layer of indexing those double checks the association of data with meta-data without even storing the original data in its buffer. This results in significant rise of accuracy even in presence of error-prone data.

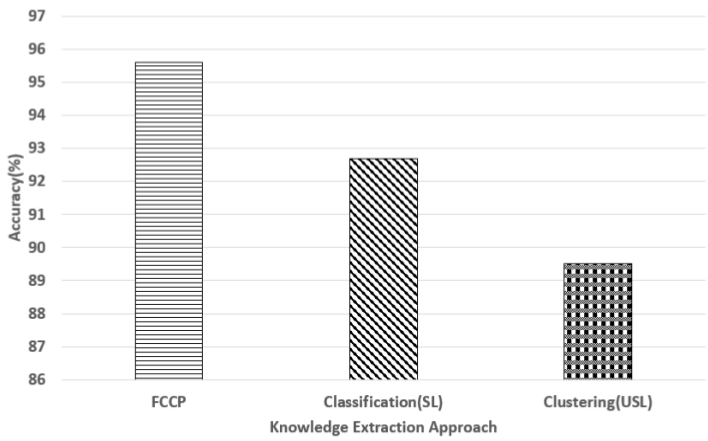

Figure 5. Comparative analysis of accuracy

3.2. Discussion towards degree of data quality performance

The proposed assessment considers the term data quality as purity of data which is defined by presence of valid values (for each respective identifier) in each cell. In simpler form, it will mean that if out of 10 rows, two rows are found to possess error prone data, than the data quality is quantitatively represented as 90% where 10% of data is found to be invalid. Hence, the prime motive of this assessment is to check the performance of both the algorithms presened by FCCP as well as existing system too on same test-environment and same data. The assessment outcome is exhibited in Figure 6.

Figure 6 states that proposed FCCP offers approximately 3.4% of improved data quality in contrast to existing schemes of SL and USL. It can be noted that USL approach still doesn't offer better data quality although it is characterized by highly scalable and interpretable charecteristics. It also bears potential to use an appropriate distance or similarity measures in order to compare the document. However, it is highly sensitive towards data preprocessing operation especially in presence of outliers which reduces data quality. It will mean that USL approach is not capable enough to find a precise alternative of the error-prone data. On the other hand, SL approach uses regularization in order to resist the possibility of overfitting and yets its capability is essentially limited to binary classification only. However, proposed scheme presents a mechanism to find the error-prone data followed by statistical correlation-based analysis in order to find the most suitable value that can be substituted with the error prone data. Hence, the quality of data is always maintained to highest possible degree irrespective of any condition in proposed scheme.

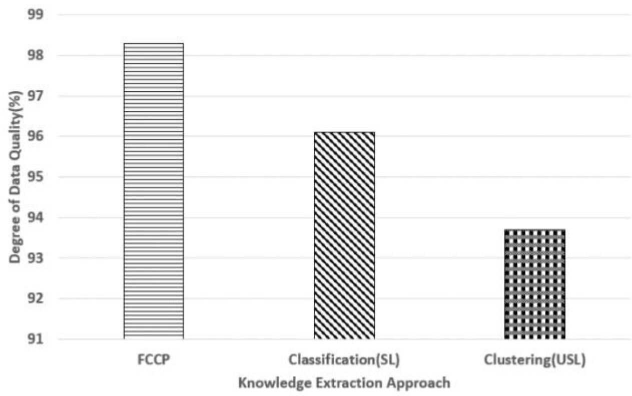

Figure 6. Comparative analysis of degree of data quality

3.3. Discussion towards preprocessing time

The prime novelty of the proposed scheme is its mechanism to perform preprocessing of the data which not only streamline the massive incoming data but also offers a significant transformation of the data. The overall duration involved in execution of the first algorithm discussed in prior section is computed in order to obtain this performance parameter. The prime reason for choosing this as performance parameter in evaluation process is because it is a novel set of operation which is responsible for tuning the data from its unarranged from to highly arranged form thereby making the data suitable for knowledge extraction process. This operation is entirely different from any existing approaches towards contextual consulting services reported in literatures.

The outcome exhibited in Figure 7 showcases that proposed FCCP is observed to offer approximatelt 2.19% of reduced preprocessing time in comparison to existing SL and USL approaches. Although, this value numerically seems to be quite low, but yet it has higher significance due to following rationale viz. Existing system of SL and USL are used for performing iterative training operation and hence the duration spend by SL (t=1.753s) and USL (t=3.882s) are basically representative of multiple steps of feature extraction and training operation in order to carry out knowledge extraction, where proposed FCCP (t=0.619s) uses this timing to perform maximum of operation that includes incoming stream acquisition, indexing the data, splitting the data for permanent and temporary buffer, and re-organization of values of respective identifier that leads to finally a highly arranged data suitable for knowledge extraction. Hence, the outcome of preprocessing time is highly significant contribution in contrast to existing schemes.

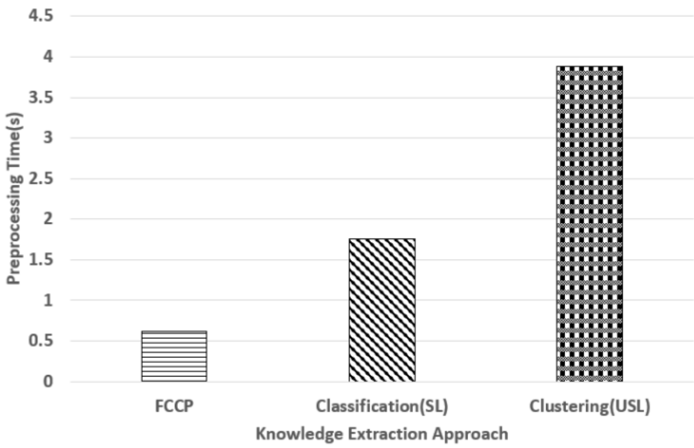

Figure 7. Comparative analysis of preprocessing time

3.4. Discussion towards memory utilization

Memory utilization is computed as total amount of memory required to carry out the complete processing of proposed scheme. Numerically, it is computed by summing up storage memory (where the final knowledge is retained) with temporary memory (where the complete analytical computation is carried out). As the proposed FCCP introduces a novel scheme of intregrating data preprocessing and optimizing data quality for knowledge discovery where specific memory units are allocated with meta-data, index values, intemediete processed data, and final data, it is essential to understand the actual data.

A closer look at Figure 8 shows that proposed FCCP offers approximately 51% of reduced memory consumption in comparison to existing SL and USL based approaches. The prime rationale behind this outcome is as follows: it can be noted that there is no significance difference between the outcome of both SL and USL approach, while USL approach is witnessed to incur slightly more consumption of memory compared to SL approach. It is known that USL offers a faster convergence which is quite appropriate measure towards exploratory data analysis and knowledge discovery however it is less effective towards capturing shapes of complex clusters that leads to failures of establishing a non-linear relationship within incoming consulting data. This phenomenon leads to retention of excessive amount of data along with its extracted knowledge (with significantly less accuracy) leading to excessive memory consumption. This problem is somewhat solved by using SL approach as it can actually deal with extensive dataset streamed from consulting origin nodes; however, Sl approach is not meant for mitigating noisy data or error-prone data which also affects its parameter tuning performance. This process leads to reterntion of knowledge along with all the intermediate processing information in the final storage units of cloud increasing memory consumption gradually. Finally, proposed scheme of FCCP works completely in different mode where the data is split into permanent and temporary memory system leading to storage of non-repeating identifiers and only the finally extracted knowedge of consulting contextual data without involving any internal training or iterative operation. Hence, significance amount of processing memory (in temporary memory system) as well as final cloud server memory system is saved in this process.

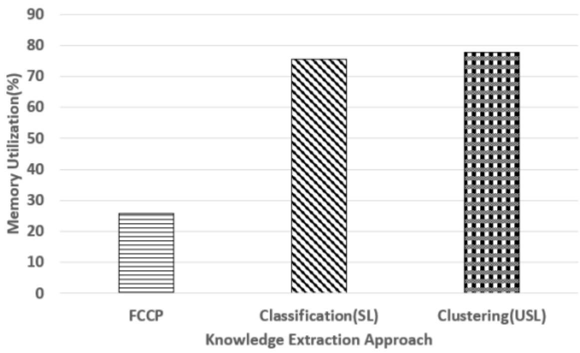

Figure 8. Comparative analysis of memory utilization

4. CONCLUSION

This paper presents a discussion of a novel framework towards contextual consulting practices in adherence to industry 4.0 automation standard. The justification behind the incorporation of industry 4.0 standard in proposed scheme can be stated as: the novelty of proposed FCCP is that incorporates the concept of data-driven decision making where a simplified knowledge delivery system is mechanism that can perform a series of transformation of data leading to final discovery of knowledge. The complete process performs this task by indexing data, splitting of data, and transformation process involved in preprocessing, followed by optimizing data quality leading to a proper data-driven decision-making approach of industry 4.0 standard. Another novelty of proposed scheme is towards incorporating the concept of decentralization scheme by developing a multiple origination points of consulting data, splitted data stored in distributed cloud servers indexed properly to establish the relationship of data in same and different storage location, and finally the ultimate outcome of knowledge discovery being stored in distributed data with another set of new indexes. It should be noted that raw data is actually stored locally while the knowledge-based data is stored in cloud servers along with its meta-data leading to a proper decentralized scheme of industry 4.0 standards. Further final outcome of the study model is proven to offer better performance with respect to multiple and diverse assessment matrix exhibiting FCCP to excel better analytical performance in contrast to existing sophisticated and iterative approaches.

REFERENCES

[1] A.-T. Järvenpää, A. Pavlik, and T. K. Gustavsson, “Contextual communicative competence in multinational infrastructure projects,” *Buildings*, vol. 11, no. 9, pp. 1–14, 2021, doi: 10.3390/buildings11090403.

[2] H. U. R. Khan, C. K. Lim, M. F. Ahmed, K. L. Tan, and M. B. Mokhtar, “Systematic review of contextual suggestion and recommendation systems for sustainable e-tourism,” *Sustainability*, vol. 13, no. 15, pp. 1–27, 2021, doi: 10.3390/su13158141.

[3] X. You, “Applying design thinking for business model innovation,” *Journal of Innovation and Entrepreneurship*, vol. 11, no. 1, pp. 1–25, 2022, doi: 10.1186/s13731-022-00251-2.

[4] M. Chiny, M. Chihab, A. A. Lahcen, O. Bencharef, and Y. Chihab, “Effect of word embedding vector dimensionality on sentiment analysis through short and long texts,” *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 12, no. 2, p. 823, Jun. 2023, doi: 10.11591/ijai.v12.i2.pp823-830.

[5] S. Rajagopal, A. Kadan, M. G. Prappanadan, and L. V. Lakshmanan, “Online news popularity prediction before publication: effect of readability, emotion, psycholinguistics features,” *IAES International Journal of Artificial Intelligence*, vol. 11, no. 2, pp. 539–545, 2022, doi: 10.11591/ijai.v11.i2.pp539-545.

[6] M. A. Lopez and J. Campbell, “Developing a communication curriculum for primary and consulting services,” *Medical Education Online*, vol. 25, no. 1, pp. 149–160, Jan. 2020, doi: 10.1080/10872981.2020.1794341.

[7] Á. N. D. S. Mata *et al.*, “Training in communication skills for self-efficacy of health professionals: a systematic review,” *Human Resources for Health*, vol. 19, no. 1, p. 30, Dec. 2021, doi: 10.1186/s12960-021-00574-3.

[8] T. N. Hindi and A. Frenkel, “The contribution of collaboration to the development of sustainable innovation in high-tech companies,” *Journal of Innovation and Entrepreneurship*, vol. 11, no. 1, pp. 1–20, 2022, doi: 10.1186/s13731-022-00259-8.

[9] J. O. Hernández and P. F. I. Casas, “Business intelligence’s self-service tools evaluation,” *Technologies*, vol. 10, no. 4, pp. 1–38, 2022, doi: 10.3390/technologies10040092.

[10] T. J. Ma, R. J. Garcia, F. Danford, L. Patrizi, J. Galasso, and J. Loyd, “Big data actionable intelligence architecture,” *Journal of Big Data*, vol. 7, no. 1, pp. 1–19, 2020, doi: 10.1186/s40537-020-00378-7.

[11] Z. MajlesiRad and A. H. H. P. Shoushtari, “Analysis of the impact of social network sites and eWOM marketing, considering the reinforcing dimensions of the concept of luxury, on tendency toward luxury brand,” *Future Business Journal*, vol. 6, no. 1, pp. 1–19, 2020, doi: 10.1186/s43093-020-00025-w.

[12] A. Renz and R. Hilbig, “Prerequisites for artificial intelligence in further education: identification of drivers, barriers, and business models of educational technology companies,” *International Journal of Educational Technology in Higher Education*, vol. 17, no. 1, pp. 1–21, 2020, doi: 10.1186/s41239-020-00193-3.

[13] M. Lechner, “Causal machine learning and its use for public policy,” *Swiss Journal of Economics and Statistics*, vol. 159, no. 1, pp. 1–15, 2023, doi: 10.1186/s41937-023-00113-y.

[14] T. Kovács, A. Ko, and A. Asemi, “Exploration of the investment patterns of potential retail banking customers using two-stage cluster analysis,” *Journal of Big Data*, vol. 8, no. 1, pp. 1–25, 2021, doi: 10.1186/s40537-021-00529-4.

[15] W. Rösler *et al.*, “An overview and a roadmap for artificial intelligence in hematology and oncology,” *Journal of Cancer Research and Clinical Oncology*, vol. 149, no. 10, pp. 7997–8006, 2023, doi: 10.1007/s00432-023-04667-5.

[16] R. Makhlouf, “Cloudy transaction costs: a dive into cloud computing economics,” *Journal of Cloud Computing*, vol. 9, no. 1, pp. 1–11, 2020, doi: 10.1186/s13677-019-0149-4.

[17] S. Nadal, P. Jovanovic, B. Bilalli, and O. Romero, “Operationalizing and automating data governance,” *Journal of Big Data*, vol. 9, no. 1, pp. 1–31, 2022, doi: 10.1186/s40537-022-00673-5.

[18] T. S. Thorisdottir and L. Johannsdottir, “Sustainability within fashion business models: a systematic literature review,” *Sustainability*, vol. 11, no. 8, pp. 1–26, 2019, doi: 10.3390/su11082233.

[19] V. Karyukin *et al.*, “On the development of an information system for monitoring user opinion and its role for the public,” *Journal of Big Data*, vol. 9, no. 1, pp. 1–45, 2022, doi: 10.1186/s40537-022-00660-w.

[20] M. Leal, F. Pisani, and M. Endler, “A blockchain-based service for inviolable presence registration of mobile entities,” *Journal of the Brazilian Computer Society*, vol. 27, no. 1, pp. 1–15, 2021, doi: 10.1186/s13173-021-00104-y.

[21] L. Lu and M. Li, “Development of a virtual interactive system for Dahua Lou loom based on knowledge ontology-driven technology,” *Heritage Science*, vol. 11, no. 1, pp. 1–14, 2023, doi: 10.1186/s40494-023-01027-x.

[22] F. Reinwald *et al.*, “Multi-level toolset for steering urban green infrastructure to support the development of climate-proofed cities,” *Sustainability*, vol. 13, no. 21, pp. 1–23, 2021, doi: 10.3390/su132112111.

[23] A. M. Ahmed, C. N. Mohammed, and A. M. Ahmad, “Web-based payroll management system: design, implementation, and evaluation,” *Journal of Electrical Systems and Information Technology*, vol. 10, no. 1, pp. 1–16, 2023, doi: 10.1186/s43067-023-00082-5.

[24] I. S. Khan, O. Kauppila, N. Fatima, and J. Majava, “Stakeholder interdependencies in a collaborative innovation project,” *Journal of Innovation and Entrepreneurship*, vol. 11, no. 1, pp. 1–17, 2022, doi: 10.1186/s13731-022-00229-0.

[25] D. H. Alahmadi and A. A. Jamjoom, “Decision support system for handling control decisions and decision-maker related to supply chain,” *Journal of Big Data*, vol. 9, no. 1, pp. 1–14, 2022, doi: 10.1186/s40537-022-00653-9.

[26] B. E. Jin and D. C. Shin, “The power of 4th industrial revolution in the fashion industry: what, why, and how has the industry changed?,” *Fashion and Textiles*, vol. 8, no. 1, pp. 1–25, 2021, doi: 10.1186/s40691-021-00259-4.

[27] X. Sun, H. Yu, W. D. Solvang, Y. Wang, and K. Wang, “The application of Industry 4.0 technologies in sustainable logistics: a systematic literature review (2012–2020) to explore future research opportunities,” *Environmental Science and Pollution Research*, vol. 29, no. 7, pp. 9560–9591, 2022, doi: 10.1007/s11356-021-17693-y.

[28] J. V. -Díez, M. Molina, J. O. -Meré, S. Sun, D. Schmidt, and W. Wellbrock, “Geometric deep lean learning: deep learning in industry 4.0 cyber–physical complex networks,” *Sensors*, vol. 20, no. 3, pp. 1–16, 2020, doi: 10.3390/s20030763.

[29] T. Sato, H. Takeuchi, S. Kondoh, and Y. Umeda, “Analyzing consultancy on production systems based on the digital triplet concept,” *Machines*, vol. 11, no. 7, pp. 1–18, 2023, doi: 10.3390/machines11070706.

[30] J. Pap, C. Mako, M. Illessy, N. Kis, and A. Mosavi, “Modeling organizational performance with machine learning,” *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 8, no. 4, p. 177, Dec. 2022, doi: 10.3390/joitmc8040177.

[31] S. A. Alanazi, M. Alruwaili, F. Ahmad, A. Alaerjan, and N. Alshammari, “Estimation of organizational competitiveness by a hybrid of one-dimensional convolutional neural networks and self-organizing maps using physiological signals for emotional analysis of employees,” *Sensors*, vol. 21, no. 11, pp. 1–29, 2021, doi: 10.3390/s21113760.

[32] M. Nurek and R. Michalski, “Combining machine learning and social network analysis to reveal the organizational structures,” *Applied Sciences*, vol. 10, no. 5, pp. 1–43, 2020, doi: 10.3390/app10051699.

[33] M. Anshari, M. Syafrudin, A. Tan, N. L. Fitriyani, and Y. Alas, “Optimisation of knowledge management (KM) with machine learning (ML) enabled,” *Information*, vol. 14, no. 1, pp. 1–15, 2023, doi: 10.3390/info14010035.

[34] O. B. Ansari, "Geo-marketing segmentation with deep learning," *Businesses*, vol. 1, no. 1, pp. 51–71, 2021, doi: 10.3390/businesses1010005.
[35] D. Cetindamar, M. Katic, S. Burdon, and A. Gunsel, "The interplay among organisational learning culture, agility, growth, and big data capabilities," *Sustainability*, vol. 13, no. 23, pp. 1–20, 2021, doi: 10.3390/su132313024.
[36] A. B. Cardoso, B. Martins, and J. Estima, "A novel deep learning approach using contextual embeddings for toponym resolution," *ISPRS International Journal of Geo-Information*, vol. 11, no. 1, pp. 1–20, 2022, doi: 10.3390/ijgi11010028.
[37] E. Camilleri and S. J. Miah, "Evaluating latent content within unstructured text: an analytical methodology based on a temporal network of associated topics," *Journal of Big Data*, vol. 8, no. 1, pp. 1–37, 2021, doi: 10.1186/s40537-021-00511-0.
[38] A. A. Funkner, M. P. Egorov, S. A. Fokin, G. M. Orlov, and S. V. Kovalchuk, "Citywide quality of health information system through text mining of electronic health records," *Applied Network Science*, vol. 6, no. 1, pp. 1–21, 2021, doi: 10.1007/s41109-021-00395-2.
[39] M. Q. Shabbir and S. B. W. Gardezi, "Application of big data analytics and organizational performance: the mediating role of knowledge management practices," *Journal of Big Data*, vol. 7, no. 1, pp. 1–17, 2020, doi: 10.1186/s40537-020-00317-6.
[40] A. Almansour, R. Alotaibi, and H. Alharbi, "Text-rating review discrepancy (TRRD): an integrative review and implications for research," *Future Business Journal*, vol. 8, no. 1, pp. 1–15, 2022, doi: 10.1186/s43093-022-00114-y.
[41] A. Cakir, Ö. Akın, H. F. Deniz, and A. Yılmaz, "Enabling real time big data solutions for manufacturing at scale," *Journal of Big Data*, vol. 9, no. 1, pp. 1–24, 2022, doi: 10.1186/s40537-022-00672-6.
[42] X. Meng, S. Li, M. M. Malik, and Q. Umer, "Machine-learning-based suitability prediction for mobile applications for kids," *Sustainability*, vol. 14, no. 19, pp. 1–14, 2022, doi: 10.3390/su141912400.

BIOGRAPHIES OF AUTHORS

Vijay Kumar Pandey holds M.CA. degree from Bangalore University in 2007. He also received B.CA. degree from Bangalore University, Karnataka in 2004. He has 7 years teaching experience. He is currently working as a full-time research scholar in Indira Gandhi National Tribal University (IGNTU), Amarkantak. He has published two research paper in international journal and conference. He can be contacted at email: vijay.research4@gmail.com.

Dr. Neeraj Kumar Rathore is an esteemed academic who assumed the position of associate professor & head in the Department of Computer Science IGNTU-A Central University, located in Amarkantak, Madhya Pradesh, India, in March 2021. Prior to this appointment, he accumulated more than fifteen years of teaching experience at renowned institutions such as Sri G.S. Institute of Technology & Science in Indore, Madhya Pradesh, India, Jaypee University of Engineering & Technology in Guna (M.P.), and Modi Institute of Technology in Rajasthan. In addition to his academic endeavors, he also gained valuable industrial experience in the IT Industry while serving as a software engineer at Computer Sciences Corporation (CSC). Academic credentials showcase his dedication to the field of Computer Science. He holds a Ph.D. in Computer Science, with a specialization in grid computing, which he earned in 2014 from Thapar University, Punjab. Furthermore, he obtained his M.E. in Computer Engineering in 2008 from the same institution and his B.E. in Computer Science and Engineering in 2006 from Rajasthan University, Rajasthan. His expertise spans various domains within the realm of computer science, including machine learning, parallel and distributed computing, grid computing, cloud computing, big data, IoT, DBMS, and data structures. He has an impressive track record of scholarly contributions, with over 72 publications to his name, encompassing 15 SCI Papers, 7 Patents, and 7 Books, all of which have appeared in esteemed international journals, conferences, and reputed books. He can be contacted at email: neeraj.rathore@igntu.ac.in.

Dr. Narayan P. Bhosale is working as an assistance professor in the Department of Computer Science at IGNTU-A Central University, located in Amarkantak, Madhya Pradesh, India. He is having more than 10 years experience in teaching and R&D. His expertise spans various domains within the realm of computer science, including cyber forensics and cyber security-computing security, digital image processing, information system and knowledge management, artificial intelligence, machine learning, blockchain technology, IoT, and data science. He can be contacted at email: narayanbhosale@igntu.ac.in.