Online ISSN : 0975-4172
Print ISSN : 0975-4350
DOI : 10.17406/GJCST

Global Journal

of Computer Science and Technology: C

Software & Data Engineering

Concept Drift Detection in Data

System Planning in Alignment

Highlights

Measurement and Classification

Potential of Big Data Analytics

Discovering Thoughts, Inventing Future

VOLUME 17 ISSUE 2 VERSION 1.0

Global Journal of Computer Science and Technology: C
Software & Data Engineering

Global Journal of Computer Science and Technology: C
Software & Data Engineering

Volume 17 Issue 2 (Ver. 1.0)

Open Association of Research Society

Incorporation No.: 0423089
License No.: 42125/022010/1186
Registration No.: 430374
Import-Export Code: 1109007027
Employer Identification Number (EIN):
USA Tax ID: 98-0673427

Global Journals Inc.

*(A Delaware USA Incorporation with "Good Standing"; **Reg. Number: 0423089**)*

Sponsors: *Open Association of Research Society*
Open Scientific Standards

Publisher's Headquarters office

Global Journals® Headquarters
945th Concord Streets,
Framingham Massachusetts Pin: 01701,
United States of America
USA Toll Free: +001-888-839-7392
USA Toll Free Fax: +001-888-839-7392

Offset Typesetting

Global Journals Incorporated
2nd, Lansdowne, Lansdowne Rd., Croydon-Surrey,
Pin: CR9 2ER, United Kingdom

Packaging & Continental Dispatching

Global Journals
E-3130 Sudama Nagar, Near Gopur Square,
Indore, M.P., Pin: 452009, India

Find a correspondence nodal officer near you

To find nodal officer of your country, please email us at *local@globaljournals.org*

eContacts

Press Inquiries: *press@globaljournals.org*
Investor Inquiries: *investors@globaljournals.org*
Technical Support: *technology@globaljournals.org*
Media & Releases: *media@globaljournals.org*

Pricing (Including by Air Parcel Charges):

For Authors:
22 USD (B/W) & 50 USD (Color)
Yearly Subscription (Personal & Institutional):
200 USD (B/W) & 250 USD (Color)

Editorial Board

Global Journal of Computer Science and Technology

Contents of the Issue

Global Journal of Computer Science and Technology: C
Software & Data Engineering
Volume 17 Issue 2 Version 1.0 Year 2017
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Concept Drift Detection in Data Stream Mining: The Review of Contemporary Literature

By B. Ramakrishna & Dr. S Krishna Mohan Rao
JNTU

Abstract- Mining process such as classification, clustering of progressive or dynamic data is a critical objective of the information retrieval and knowledge discovery; in particular, it is more sensitive in data stream mining models due to the possibility of significant change in the type and dimensionality of the data over a period. The influence of these changes over the mining process termed as concept drift. The concept drift that depict often in streaming data causes unbalanced performance of the mining models adapted. Hence, it is obvious to boost the mining models to predict and analyse the concept drift to achieve the performance at par best. The contemporary literature evinced significant contributions to handle the concept drift, which fall in to supervised, unsupervised learning, and statistical assessment approaches. This manuscript contributes the detailed review of the contemporary concept-drift detection models depicted in recent literature. The contribution of the manuscript includes the nomenclature of the concept drift models and their impact of imbalanced data tuples.

Keywords: concept drift, change point, data stream mining, ensemble classifiers, class imbalance, misclassification, supervised and unsupervised learning.

GJCST-C Classification: C.1.1, H.2.8

Strictly as per the compliance and regulations of:

Concept Drift Detection in Data Stream Mining: The Review of Contemporary Literature

B. Ramakrishna [α] & Dr. S Krishna Mohan Rao [σ]

Abstract- Mining process such as classification, clustering of progressive or dynamic data is a critical objective of the information retrieval and knowledge discovery; in particular, it is more sensitive in data stream mining models due to the possibility of significant change in the type and dimensionality of the data over a period. The influence of these changes over the mining process termed as concept drift. The concept drift that depict often in streaming data causes unbalanced performance of the mining models adapted. Hence, it is obvious to boost the mining models to predict and analyse the concept drift to achieve the performance at par best. The contemporary literature evinced significant contributions to handle the concept drift, which fall in to supervised, unsupervised learning, and statistical assessment approaches. This manuscript contributes the detailed review of the contemporary concept-drift detection models depicted in recent literature. The contribution of the manuscript includes the nomenclature of the concept drift models and their impact of imbalanced data tuples.

Keywords: concept drift, change point, data stream mining, ensemble classifiers, class imbalance, misclassification, supervised and unsupervised learning.

I. Introduction

Rapid expansion of Information and Communication Technologies (ICT) has led to exponential growth in the volume of data generated. According to a survey from IDC, in the near future the quantum of data that is generated will run to trillions of gigabytes [1]. It is imperative that there is need for robust tools and solutions for handling the huge volume of data generated from varied range of applications. Eventually the huge volumes of data that are generated demand more effective techniques of data management.

Data mining is one of the major process in the data management. Profoundly the data mining solutions are about initially gathering data and processing them in offline mode. Predictive models are usually trained on the basis of historical data provided as pair of data (input and output). The trained models are used for prediction of output for new unseen input data.

Streaming data can't be processed simultaneously due to the quantum of data that is generated regularly. It is highly complex to accommodate the data in to machine's main memory and the online processing of data is the only right method that could be adapted. Predictive models can be trained either in an incremental manner by continuous update or by ensuring retaining of the model using batches of data.

In the constantly changing environments, the data distribution might change over course of time and it could lead to conditions of concept drift [2], [3]. Concept drift is the changes in the conditional distribution of varied output (for example, the target variable for the input features) despite of the input remaining unchanged.

A classic example of real concept drift is about the change in user level interests whilst following an online news stream. For instance, though the distribution of a news documents that are often relayed might remain the same, still the conditional distribution of interesting news documents for the user might undergo changes. The process of adaptive learning reflects upon the predictive models online while their operation might be responding to concept drifts.

Phenomenal developments have taken place in terms of concept drift and there were many drift-aware adaptive learning algorithms were developed. The scope of the problem being very wide and spans over varied topics, not much of comprehensive survey is envisaged. Though the concept is relatively new, still there was some kind of adaptive learning algorithms that were proposed earlier.

Considering the quantum of develop ments that has taken place in the subject of concept drift, this paper focus on comprehensive summary of research done for gaining insights in to concepts unification and terminology and also to survey of contemporary methodologies and techniques that are investigated in the past.

II. The Problem Description

a) Misclassification

Profoundly, in the case of misclassification, the minority class is highly complex than the majority class. For instance, the spam class in the spam filters and the fraud class in a credit card application. Hence, misclassifying-class example is highly a costlier factors. Predominantly the performance of many of the conventional machine learning algorithms compromise

Author α: Assistant professor, Sphoorthy Engineering College, Hyderabad. e-mail: ramakrishnabasude@gmail.com
Author σ: Professor Gandhi Institute for Technology, Bhubaneswar. e-mail: krishnamohan6@yahoo.com

the facets of misclassification, as they presume only balanced class distribution. Training procedure with target of maximizing overall accuracy usually leads to higher levels of probability of induced classifier that predicts as the instance of majority class along with low It is imperative to envisage majority class has higher accuracy levels while the minority class has much lower accuracy amidst ranging between 0-10% [4]. Misclassification resulting because of imbalance due to classifiers like the decision trees [5], KNN (K-Nearest Neighbour) [6], [7], [8], Neural Networks [9] and SVM [10], [11] that were reviewed. Classifiers offer a balanced degree of predictive performance over all classes that are required.

Profoundly the percentage of minority class in a data set is used in the researches to detail the level of imbalance in the data [12], list of varied illustrations in every class [13], size ratio amidst of classes [14]. Coefficient of variance is used in [15] that are less straight forward. Detailing of imbalance status might not be a crucial issue in offline learning, but it becomes more significant for online learning as there is hardly any static data over the online scenarios.

It is very essential to have some automatic evaluation for detailing the updated imbalance degree and techniques that monitor the changes over the Misclassification status. The facets of changes in the misclassification are directly coherent to concept drift.

b) Concept drift

Concept drift could take place when the joint probability P 9 (x; y) changes [16], [17]. Concept drift will manifest three fundamental forms of changes pertaining to three key variables in the Bayes' theorem [18].

1) Drift by prior probability (a change in learnt decision boundary):

The prior probability of circle class is reduced and the change can lead to misclassification. Identification of drift using the prior probability is simple and it is distance between two concepts that are estimated depending on the distance assessment methods like the Total Variation Distance and Hellinger Distance.

2) Drift by condition (decision boundary change influenced by condition):

True decision boundary remains unaffected. In the earlier research, authors have claimed that such types of drift are result of incomplete representation of true distribution over the current data that profoundly needs supplement data information for the learning model [19].

The subset of covariate attributes will have conditional probability distribution over varied possible values of covariate attributes for every specific class. Conditional drift is weighted sum of distances amidst every probability distributions from varied time period, wherein the weights are average probability of class amidst the time periods in sequence.

3) Drift by posterior probability (a change influenced by the conflict of old and new decision boundary):

True boundary amidst classes varies post drift and the earlier learnt discrimination function does not apply the changes more. In a different dimension, it can be stated that the old function becomes completely or partially unfit and the learning models are required to adapt to new dimensions.

For every subset of covariate attributes, there is a probability distribution amidst the class labels for every combination of covariate values at every time period. Hence the posterior drift can be estimated as weighted sum of distances amidst the probability distributions wherein the weighted sum of such distances amid the probability distributions wherein the weights are average probability over two periods of specific value to an covariate attribute subset.

Posterior distribution changes signify the fundamental changes among the data generation function, which is classified as a true concept drift. The other two types relate to virtual concept drift [20] that do not change the decision boundaries. In real-term conditions, one type of concept drift might appear with combination with other kind of concept drifts.

III. Review of State of The Art Models

In this section, the contemporary models pertaining to concept drift detection of streaming data mining from the contemporary literature. Overall models that were reviewed in varied context and it is a join detection of concept drift and misclassification. Concept drift detection using incremental learning and concept drift detection based on statistical methods.

a) Joint detection of Concept Drift and Misclassification

Few of the researches have made effort to address the joint problem of concept drift and misclassification, because of rising need from practical problems [21], [22].

Uncorrelated Bagging is one of the old algorithm that is built to ensemble classifiers that are trained over a more balanced set of data based on re sampling and overcome the concept drift passively by weighing the base classifier having discriminative power [23], [24], [25].

Selectively the recursive approaches like REA [26] and SERA [27] adapt same kind of ideas for Uncorrelated Bagging of developing an ensemble weighted classifiers but with a smarter level of oversampling technique. Learn ++NIE and the Learn++CDS are some of the contemporary algorithms that tackle misclassification based on oversampling technique SMOTE [13] or sub-ensemble technique and

the concept drift based on a dynamic weighting strategy [28].

HUWRS.IP [29] develops HUWRS [30] to handle the imbalanced data streams by focusing on instance propagation scheme that relies on Naïve Bayes classifier and it uses Hellinger distance as a major weighting measure for concept drift detection.

All such approached relate to chunk-based learning algorithms, and the core techniques work over a batch of data that is received at every step. It is very complex to develop a true online algorithm for concept drift due to the issues like measuring minority class statistics based on one illustration at a time [31].

In order to handle the misclassification and concept drift in the form of an online fashion, some of the methods were proposed in the recent past. DDM – OCI [32] is among one of the contemporary algorithms that were proposed for detection of concept drift actively over the imbalanced data streams online. It tracks the reduction over the minority-class recall and upon observing any kind of significant drop, the drift shall be reported. The solution was very effective when minority-class recall is impacted by concept drift but when majority class might be adversely impacted.

LFR (Linear Four Rates) is the other approach proposed for improving the DDM-OCI that monitors four rates based on confusion matrix- wherein the minority class recall and the precision, majority-class recall and precision that is statistically-supported bounds towards any kind of drift detection [33]. If any of the four rates if found exceeding the bound, the drift shall be confirmed. In PAUC (Prudential AUC) [34], [35] the emphasis is on developing an overall performance measure for online scenarios, and used as concept drift indicator. But accessibility to the historical data is imperative for the system. DDM-OCI, PAUC and LFR are very active drift detectors that are designed for imbalanced data streams and they are independent over the classification algorithms. Such significant constraint of such models is reset of the learning process if concept drift is considered. It could be infeasible in terms of handling the mis classification.

In addition to the above set of approaches, perceptr on oriented algorithms like ESOS-ELM [36], RLSACP [37] and ONN [38] focus on the classification model for non-stationery environments in a passive manner and comprise mechanisms to address misclassification. RLSACP and ONN are some of the single model approaches comprising similar set of modelling and framework.

CID (Class Imbalance Detection) approach was proposed with a varied objective towards concept drift [39]. For defining the imbalanced degree that is suitable for online learning, a real-time indicator was proposed which is based on time decayed class size, the size pertaining to every class in the data stream. It is updated incrementally at every time using the time decay factor that emphasizes current status of data and it weakens the effect of old data. Any kind of current imbalance status is reported and it provides information pertaining to which classes belong to minority and majority classes.

b) Concept Drift Detection by Incremental learning

Incremental learning is a new dimension in which the concept drift is identified with. Many of the models that were proposed earlier focused on incremental learning wherein the historical models were considered for forming the ensembles. Following are some of the contemporary incremental learning models SEA (Streaming Ensemble Algorithm) [40] uses simple majority voting, the DIC (Dynamic Integration of Classifiers) approach [41] combines historical models with novel model of data training using the dynamic selection (DS), DVS (Dynamic Voting with Selection) and Dynamic Voting (DV).

The other benchmark called AUE2 (Accuracy Updated Ensemble) [42] adapts weighted voting as a combination scheme, where the weights that assigned to individual models are defined in terms of mean squared errors of the models. Learn++ algorithm [43] unlike DIC and AUE2 weighs on the current performance over the Non-Stationary Environments, which assign weights to varied individual models depending on the current and the earlier data.

The model discussed in [44] reflects upon Inductive Transfer (TIX) approach which works on varied methods to gain insights to historical models like given a new chunk of data and the outputs of historical models over training data that are used as features of training data, and a new model is developed with augmented training data. In the instance of building linear models that are built on learning process, TIX can be perceived as one of the special weighted voting scheme, a linear combination of original features of training data can be perceived as outcome of a linear model based on the training data that is original.

The other ensemble model DDD (Diversity for Dealing with Drifts) method is discussed in [16]. The method focus on using the ensembles as single model for a chosen time step.

Existence of concept drift leads to various models with positive and negative effects in terms of learning the current concept. It is very important that whilst getting the positive effects, preventing the negative impacts is also very important. Preserving historical models induce overheads for both storage and computation. Such issues are not usually addressed in DIC, TIX, Learn++NSE. Despite that DS/DVS scheme of DIC and time-adjust error schemes of Learn++NSE shall be resourceful for choosing historical models for preserving, and such adaptations need not be evaluated.

SEA and AUE2 usually control the number of preserved models in the conditions of a predefined threshold. Both SEA and AUE2 assess the quality of individual models based on accuracy perspective. Major difference in the way SEA and AUE2 assess is that, SEA takes to account overall accuracy of ensembles of current training data, but AUE2 takes in to account every individual model in consideration over the training data in direct manner.

Not many of the existing methods of ensemble that are used for incremental learning has focused on ensemble diversity in an explicit manner, though the diversity is considered to play a critical role in ensembles [45], [46].

c) Concept Drift Detection by Statistical Measures

In [33], the related statistical change detection model was proposed to handle the imbalanced data streaming, wherein the proposed model monitor multiple performance metrics. The technique monitors true positive rate and false positive rate, the true negative rates and the false positive rates attained from the confused matrix of the classification. Unlike the traditional matrix that reflects a biased majority class, the confusion matrix depicts more detailed view that is essential for addressing imbalance class problems.

In [47], the incremental model was proposed in which the EWMA was proposed to signal deviation in average error rate, by considering the quantum of standard deviations from the mean.

DOD (Degree of Drift) which is an window-based model identifies the drifts by computing the distance map of all the samples over the current chunk and its nearest neighbours from the earlier chunk [48]. The DOD is computed based on distance map and if it increases by a parameter, the drift is signalled.

In [49], the Paired Learners approach is proposed which adapts a pair of reactive learner which is trained based on the chunk of data. The model is a stable learner and trained over all the earlier data. The variation of accuracy amidst the two approaches depicts the drift. The disagreement is captured over binary value circular list. Also, the increase in the quantum of ones that are beyond change threshold is signalled as a drift, which is managed by replacing the stable model with a reactive one.

In [50], a contemporary model was proposed that depends on the observation of randomly chosen training and testing the samples from a chunk of data which should lead to good accuracy of prediction, unless the window have any kind of non-stationary data. The usually adapted model of classifier's cross validation evaluation [51] is the fundamental for the aforesaid model.

The OLINDDA (Online Novelty and Drift Detection Algorithm) adapts the K-means data clustering for monitoring continuously and adapt to the emerging data [52]. The short term memory queue holds the unknown samples and they are clustered periodically and merged to existing similar cluster profiles or the modern profile to the pol of clusters.

In [53], the MINAS were proposed which relies on micro clusters to obtain incremental stream clustering algorithm. It is an extended model to OLINLINDAs approach used for multi class problem.

Some of the similar techniques that rely on clustering for defining the boundaries of the known data are Woo Ensemble [54], ECS Miner [55], and DETECTNOD algorithm [56]. Samples falling out from the clusters are treated as suspicious [54] [56] or the other way as outliers [55]. The difference or similarity amidst the defined clusters and suspicious samples are estimated on the account of density that is observed. If similarity attains the suspicious or outlier samples that are incorporated towards corresponding clusters, it concludes the concept drift.

In [57], the GC3 approach is an improvisation with a grid density oriented clustering algorithm, wherein the novelty is estimated by considering newly appearing grids in the data space. Such methods [52], [54],[53], [56], [55], [57] face challenge of curse of dimensionality and issues of distance methods detection of concept drift in the binary data spaces. Such techniques are effective for multi-class classification problems and many classes might emerge or wane during the process of stream.

In [58], the COC (Change of Concept) treats every feature as an independent stream of data and screens the correlation amidst current chunk and the training chunk that has to be referred. The change observed in the average correlation is used for signal of change. Pearson correlation is used which makes the normality assumption to a distribution.

In [59], the non-parametric unlabelled approach was proposed in the model of HDDDM (Hellinger Distance Drift Detection Methodology). Hellinger distance is used a measure to change in the distribution.

In [60] and [61], the PCA (Principal Component Analysis) is used for drift detection computationally efficient for high dimensional data streams. Such techniques reduce set of features essentially to be monitored. Both the methods are contrary in the selection of principal components. PCA models envisage issues because of considerable false alarming rate when compared to the other kind of multivariate distribution models.

In [62], [63], [64], [65] consider the classification process by taking in to account the posterior probability estimates of classifiers, for identifying the drift. It can be used with probabilistic classifiers that have output of class probabilities before thresholding them to generate any kind of final class label. By tracking the posterior probability estimates, the

intensity of drift detection task is reduced to the levels of monitoring univariate stream of values, which enable the process more computationally efficient.

Such methods are very effective in reducing the false alarm rates, but their dependence on using the probabilistic models creates implications in terms of the applicability. The methods also impact any kind of change to the posterior distribution in terms of margin samples. Changes that deviate from the margin of classifier are considerably less critical than the classification process, however, none of the approaches offer robustness against them.

d) Observations

In many of the existing studies the focus is on development of drift detection methods and techniques for addressing the real drift. There is not much of research that has taken place in the domain which might impact the classification purpose and the performance. Despite that afore discussed drifts do not impact the true decision boundaries, it can lead to a better levels of decision boundary. The current techniques for handling the real drift might not be effective for any virtual drift, however, they offer different scenarios to learn and need varied solutions. In the instance of methods to address the real drift that is chosen to reset and retrain the classifier, the old concepts are ignored and the new concepts are learnt, which might not be an appropriate strategy to be used in virtual drift.

It could be more effective for calibrating the existing classifier rather than retraining them. Also, the techniques for handling the actual drift very much depends on feedback about the classifier's performance, whilst the techniques towards handling virtual drift shall operate even without the feedback [66]. It is imperative from the above review that all the three types discussed has significance and still there is scope for improvement in the models.

e) Future Research Objectives

The future challenges in concept drift could be attributed to the scope of scalability, sturdiness and efficacy of the models right from the levels of adaptation to more interpretable solutions, which can reduce dependency over the time and improve accurate feedback. Even moving from the adaptive algorithms to adaptive systems which could impact the complete knowledge discovery process apart from automating adaptation of the decision models. Few of the challenges envisaged in the process are discussed in [67].

The outline of the key issues have to be addressed by the research to ensure that a significant progress in the area of pre-processing techniques for the data stream problems.

Limited number of online and supervised discretises that were proposed in the literature reflect that the adjustments turnout to be more abrupt. But the problem is addressed to an extent by inclusion of class information in the discrete zation process. The tweaks that are abrupt and the ones that are labelling are two of the key concerns which must be addressed in the future researches.

There is integral need for wrapper-based solutions that were not explored much in the earlier researches. Pure wrapper based on the online learning solutions could limit the computational costs because of the discriminative ability and adaptability to drift. Also, there is need for further research in terms of feature and the instance selection methods which can directly impact the problem of concept drift.

IV. Conclusion

In this paper, the categorization of the existing models of adaptive learning strategies based on the conceptual models, and the ones that are able to adapt the concept drift in addition to the contemporary techniques. Majority of the concept drift models presume that the changes taken place in the hidden contexts that are complex to be identified in the adaptive learning system. Because of the aforesaid reasons, concept drift is considered as an unpredictable and its detection is profoundly a reactive approach.

Numerous application settings wherein the concept drift is considered to be reappearing based on time line and based on varied objects over the modelled domain. The seasonal effects comprising vague periodicity towards a certain subgroup of object can be very common. Availability of external levels of contextual information or the extraction from the hidden contexts based on the predictive features might assign in handling the recurrent concept drift in an effective manner.

Majority of the earlier works on the concept drift detection models reviewed in this survey do not address the issues of representation bias which is prevalent in many of the adaptive systems that can direct a specific kind of behaviour. However, when there is any kind of reinforcement feedback or any kind of closed loop control towards learning mechanism, it can't evaluate and compare the performance of the concept drift based on the historical data. Hence, it can be stated that there is need for more emphatic studies that support in embedded concept drift handling in real operational settings towards proper kind of validation. While majority of the works towards handling concept drift has considered the supervised settings having immediate availability, still the actual problems looms much wider.

In the process of a supervised learning that emerges over data, and the case of delayed set of on-demand labelling over the supervised learning,

adaptation mechanisms are to be investigated. The related research in the domain of concept drift is much beyond the application of machine learning, pattern recognition and the data mining solutions and there could be more explorative solutions in the domain.

References Références Referencias

1. Gantz, John, and David Reinsel. "The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east." IDC iView: IDC Analyze the future 2007.2012 (2012): 1-16.
2. Schlimmer, Jeffrey C., and Richard H. Granger. "Incremental learning from noisy data." Machine learning 1.3 (1986): 317-354.
3. Widmer, Gerhard, and MiroslavKubat. "Learning in the presence of concept drift and hidden contexts." Machine learning 23.1 (1996): 69-101.
4. Kubat, Miroslav, Robert C. Holte, and Stan Matwin. "Machine learning for the detection of oil spills in satellite radar images." Machine learning 30.2-3 (1998): 195-215.
5. Japkowicz, Nathalie, and Shaju Stephen. "The class imbalance problem: A systematic study." Intelligent data analysis 6.5 (2002): 429-449.
6. Kubat, Miroslav, and Stan Matwin. "Addressing the curse of imbalanced training sets: one-sided selection." ICML. Vol. 97. 1997.
7. Batista, Gustavo EAPA, Ronaldo C. Prati, and Maria Carolina Monard. "A study of the behavior of several methods for balancing machine learning training data." ACM Sigkdd Explorations Newsletter 6.1 (2004): 20-29.
8. Mani, Inderjeet, and I. Zhang. "kNN approach to unbalanced data distributions: a case study involving information extraction." Proceedings of workshop on learning from imbalanced datasets. 2003.
9. Visa, Sofia, and AncaRalescu. "Issues in mining imbalanced data sets-a review paper." Proceedings of the sixteen midwest artificial intelligence and cognitive science conference. Vol. 2005. sn, 2005.
10. Yan, Rong, et al. "On predicting rare classes with SVM ensembles in scene classification." Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on. Vol. 3. IEEE, 2003.
11. Wu, Gang, and Edward Y. Chang. "Class-boundary alignment for imbalanced dataset learning." ICML 2003 workshop on learning from imbalanced data sets II, Washington, DC. 2003.
12. Van Hulse, Jason, Taghi M. Khoshgoftaar, and Amri Napolitano. "Experimental perspectives on learning from imbalanced data." Proceedings of the 24th international conference on Machine learning. ACM, 2007.
13. Chawla, Nitesh V., et al. "SMOTE: synthetic minority over-sampling technique." Journal of artificial intelligence research 16 (2002): 321-357.
14. López, Victoria, et al. "An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics." Information Sciences 250 (2013): 113-141.
15. Hoens, Thomas Ryan. Living in an imbalanced world. University of Notre Dame, 2012.
16. Minku, Leandro L., and Xin Yao. "DDD: A new ensemble approach for dealing with concept drift." IEEE transactions on knowledge and data engineering 24.4 (2012): 619-633.
17. Minku, Leandro L., Allan P. White, and Xin Yao. "The impact of diversity on online ensemble learning in the presence of concept drift." IEEE Transactions on Knowledge and Data Engineering 22.5 (2010): 730-742.
18. Kelly, Mark G., David J. Hand, and Niall M. Adams. "The impact of changing populations on classifier performance." Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 1999.
19. Elwell, Ryan, and RobiPolikar. "Incremental learning of concept drift in nonstationary environments." IEEE Transactions on Neural Networks 22.10 (2011): 1517-1531.
20. Hoens, T. Ryan, RobiPolikar, and Nitesh V. Chawla. "Learning from streaming data with concept drift and imbalance: an overview." Progress in Artificial Intelligence 1.1 (2012): 89-101.
21. Pan, Shirui, et al. "Graph ensemble boosting for imbalanced noisy graph stream classification." IEEE transactions on cybernetics 45.5 (2015): 954-968.
22. Sousa, Maria Rocha, João Gama, and ElísioBrandão. "A new dynamic modeling framework for credit risk assessment." Expert Systems with Applications 45 (2016): 341-351.
23. Gao, Jing, et al. "Classifying data streams with skewed class distributions and concept drifts." IEEE Internet Computing 12.6 (2008).
24. Gao, Jing, et al. "A general framework for mining concept-drifting data streams with skewed distributions." Proceedings of the 2007 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, 2007.
25. Wu, Ke, et al. "Classifying imbalanced data streams via dynamic feature group weighting with importance sampling." Proceedings of the 2014 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, 2014.
26. Chen, Sheng, and Haibo He. "Towards incremental learning of non stationary imbalanced data stream: a multiple selectively recursive approach." Evolving Systems 2.1 (2011): 35-50.

27. Chen, Sheng, and Haibo He. "Sera: selectively recursive approach towards non stationary imbalanced stream data mining." Neural Networks, 2009. IJCNN 2009. International Joint Conference on. IEEE, 2009.
28. Ditzler, Gregory, and RobiPolikar. "Incremental learning of concept drift from streaming imbalanced data." ieee transactions on knowledge and data engineering 25.10 (2013): 2283-2301.
29. Hoens, Thomas Ryan, and Nitesh V. Chawla. "Learning in non-stationary environments with class imbalance." Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2012.
30. Hoens, T. Ryan, Nitesh V. Chawla, and RobiPolikar. "Heuristic updatable weighted random subspaces for non-stationary environments." Data Mining (ICDM), 2011 IEEE 11th International Conference on. IEEE, 2011.
31. Ditzler, Gregory, et al. "Learning in nonstationary environments: A survey." IEEE Computational Intelligence Magazine 10.4 (2015): 12-25.
32. Wang, Shuo, et al. "Concept drift detection for online class imbalance learning." Neural Networks (IJCNN), The 2013 International Joint Conference on. IEEE, 2013.
33. Wang, Heng, and Zubin Abraham. "Concept drift detection for streaming data." Neural Networks (IJCNN), 2015 International Joint Conference on. IEEE, 2015.
34. Brzezinski, Dariusz, and Jerzy Stefano wski. "Prequential AUC for classifier evaluation and drift detection in evolving data streams." International Workshop on New Frontiers in Mining Complex Patterns. Springer International Publishing, 2014.
35. Brzezinski, Dariusz, and Jerzy Stefano wski. "Prequential AUC: Properties of the Area Under the ROC Curve for Data Streams with Concept Drift."
36. Mirza, Bilal, Zhiping Lin, and Nan Liu. "Ensemble of subset online sequential extreme learning machine for class imbalance and concept drift." Neuro computing 149 (2015): 316-329.
37. Ghazikhani, Adel, Reza Monsefi, and HadiSadoghiYazdi. "Recursive least square perceptron model for non-stationary and imbalanced data stream classification." Evolving Systems 4.2 (2013): 119-131.
38. Ghazi khani, Adel, Reza Monsefi, and Hadi Sadoghi Yazdi. "Online neural network model for non-stationary and imbalanced data stream classification." International Journal of Machine Learning and Cybernetics 5.1 (2014): 51-62.
39. Wang, Shuo, Leandro L. Minku, and Xin Yao. "A learning framework for online class imbalance learning." Computational Intelligence and Ensemble Learning (CIEL), 2013 IEEE Symposium on. IEEE, 2013.
40. Street, W. Nick, and Yong Seog Kim. "A streaming ensemble algorithm (SEA) for large-scale classification." Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2001.
41. Tsymbal, Alexey, et al. "Dynamic integration of classifiers for handling concept drift." Information fusion 9.1 (2008): 56-68.
42. Brzezinski, Dariusz, and Jerzy Stefanowski. "Reacting to different types of concept drift: The accuracy updated ensemble algorithm." IEEE Transactions on Neural Networks and Learning Systems 25.1 (2014): 81-94.
43. Polikar, Robi, et al. "Learn++: An incremental learning algorithm for supervised neural networks." IEEE transactions on systems, man, and cybernetics, part C (applications and reviews) 31.4 (2001): 497-508.
44. Forman, George. "Tackling concept drift by temporal inductive transfer." Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 2006.
45. Brown, Gavin, et al. "Diversity creation methods: a survey and categorisation." Information Fusion 6.1 (2005): 5-20.
46. Tang, E. Ke, Ponnuthurai N. Suganthan, and X in Yao. "An analysis of diversity measures." Machine learning 65.1 (2006): 247-271.
47. Ross, Gordon J., et al. "Exponentially weighted moving average charts for detecting concept drift." Pattern Recognition Letters 33.2 (2012): 191-198.
48. Sobhani, Parinaz, and Hamid Beigy. New drift detection method for data streams. Springer Berlin Heidelberg, 2011.
49. Bach, Stephen H., and Marcus A. Malo of. "Paired learners for concept drift." Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on. IEEE, 2008.
50. Harel, Maayan, et al. "Concept Drift Detection Through Re sampling." ICML. 2014.
51. Kohavi, Ron. "A study of cross-validation and bootstrap for accuracy estimation and model selection." Ijcai. Vol. 14. No. 2. 1995.
52. Spinosa, Eduardo J., André Ponce de Leon F de Carvalho, and João Gama. "Olindda: A cluster-based approach for detecting novelty and concept drift in data streams." Proceedings of the 2007 ACM symposium on Applied computing. ACM, 2007.
53. Faria, Elaine R., João Gama, and André CPLF Carvalho. "Novelty detection algorithm for data streams multi-class problems." Proceedings of the 28th annual ACM symposium on applied computing. ACM, 2013.
54. Ryu, Joung Woo, et al. "An efficient method of building an ensemble of classifiers in streaming

data." International Conference on Big Data Analytics. Springer Berlin Heidelberg, 2012.
55. Masud, Mohammad, et al. "Classification and novel class detection in concept-drifting data streams under time constraints." IEEE Transactions on Knowledge and Data Engineering 23.6 (2011): 859-874.
56. Hayat, MortezaZi, and Mahmoud Reza Hashemi. "A dct based approach for detecting novelty and concept drift in data streams." Soft Computing and Pattern Recognition (SoCPaR), 2010 International Conference of. IEEE, 2010.
57. Sethi, Tegjyot Singh, MehmedKantardzic, and Hanquing Hu. "A grid density based framework for classifying streaming data in the presence of concept drift." Journal of Intelligent Information Systems 46.1 (2016): 179-211.
58. Lee, Jeonghoon, and Frederic Magoules. "Detection of concept drift for learning from stream data." High Performance Computing and Communication & 2012 IEEE 9th International Conference on Embedded Software and Systems (HPCC-ICESS), 2012 IEEE 14th International Conference on. IEEE, 2012.
59. Ditzler, Gregory, and RobiPolikar. "Hellinger distance based drift detection for nonstationary environments." Computational Intelligence in Dynamic and Uncertain Environments (CIDUE), 2011 IEEE Symposium on. IEEE, 2011.
60. Kuncheva, Ludmila I., and William J. Faithfull. "PCA feature extraction for change detection in multidimensional unlabeled data." IEEE transactions on neural networks and learning systems 25.1 (2014): 69-80.
61. Qahtan, Abdulhakim A., et al. "A pca-based change detection framework for multidimensional data streams: Change detection in multidimensional data streams." Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2015.
62. Dries, Anton, and Ulrich Rückert. "Adaptive concept drift detection." Statistical Analysis and Data Mining 2.5-6 (2009): 311-327.
63. Lindstrom, Patrick, Brian Mac Namee, and Sarah Jane Delany. "Drift detection using uncertainty distribution divergence." Evolving Systems 4.1 (2013): 13-25.
64. Dredze, Mark, Tim Oates, and Christine Piatko. "We're not in Kansas anymore: detecting domain changes in streams." Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2010.
65. Žliobaite, Indre. "Change with delayed labeling: When is it detectable?." Data Mining Workshops (ICDMW), 2010 IEEE International Conference on. IEEE, 2010.
66. Gama, João, et al. "A survey on concept drift adaptation." ACM Computing Surveys (CSUR) 46.4 (2014): 44.
67. Zliobaite, Indre, et al. "Next challenges for adaptive learning systems." ACM SIGKDD Explorations Newsletter 14.1 (2012): 48-55.

Global Journal of Computer Science and Technology: C
Software & Data Engineering
Volume 17 Issue 2 Version 1.0 Year 2017
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Measurement and Classification of Network Traffic Analysis using Hardoop

By Tanmay Paul
Adamas Institute of Technology

Abstract- Network traffic can classified as a process which list computer network based on some parameters like port number and protocols into some traffic classes like undesired, sensitivity etc. Traffic can be implemented differently to differentiate the service required for the user for the specific purpose. The large demand of increase in internet users and increase in bandwidth required for various applications are escalating day by day. The traffic data needs to be classified and analyzed with certain tools. Hardoop is the tool which performs the task in very time efficient manner. Hardoop actually run on commodity hardware which process this huge data with hive. Traffic analysis, measurement and classification are done by hardoop based tools at various parameters of packet and flow level. The derived result is used by network administrator for resolving networking related issue. The measurement of internet traffic and analysis has been implemented from long before but the problem is recent years the user in internet has escalated dramatically. We proposed network traffic management system for analyzing internet traffic of multi-terabytes in extensible manner to perform HTTP, ICMP, UDP, TCP and IP.

Keywords: *network traffic, hardoop, traffic management and analysis, HDFS, HIVE, IP.*

GJCST-C Classification: *C.2.3*

Strictly as per the compliance and regulations of:

Measurement and Classification of Network Traffic Analysis using Hardoop

Tanmay Paul

Abstract- Network traffic can classified as a process which list computer network based on some parameters like port number and protocols into some traffic classes like undesired, sensitivity etc. Traffic can be implemented differently to differentiate the service required for the user for the specific purpose. The large demand of increase in internet users and increase in bandwidth required for various applications are escalating day by day. The traffic data needs to be classified and analyzed with certain tools. Hardoop is the tool which performs the task in very time efficient manner. Hardoop actually run on commodity hardware which process this huge data with hive. Traffic analysis, measurement and classification are done by hardoop based tools at various parameters of packet and flow level. The derived result is used by network administrator for resolving networking related issue. The measurement of internet traffic and analysis has been implemented from long before but the problem is recent years the user in internet has escalated dramatically. We proposed network traffic management system for analyzing internet traffic of multi-terabytes in extensible manner to perform HTTP, ICMP, UDP, TCP and IP.

Keywords: network traffic, hardoop, traffic management and analysis, HDFS, HIVE, IP.

I. Introduction

The collection of different servers, computers, peripherals, devices when connected to one another for secure mean of communication is described as network which is mainly used for sharing data, or as a means of communication. The process of monitoring network traffic involves managing and analyzing network to overcome any discrepancy that might be a problem for the network. The amount of data involve in communication between network is described as network traffic. The network packet [1] mostly comprised of network data which makes the load within the network. The monitoring mainly involves analyses incoming and outgoing packet. The measurement of traffic over a particular network is called traffic measurement. There are basically two types of techniques involved. Firstly the active techniques and the secondly is the passive technique. Active [2]are more accurate and instructive and the main drawback is that it may over crowd the network by infuse with artificial inquest traffic whereas passive [3] run on the background which can be used to implement network analyzing action and the drawback is that supervise on all network[4]. The main challenge of internet traffic measurement is firstly flow statistics computation time and secondly single node failure. To overcome this problem we implement [5] hardoop framework. Hardoop is actually an open source software framework for large set data processing and storage. It provides necessary possibilities of scaling and fault tolerance which are the most important in networking. We also implement Map Reduce model to resolve the inconsistency in between the hardoop data distribution and network monitoring where data is recorded and splinted and dispense them into cluster for individual processing. The related packets may spread across different splits, thus dislocating traffic structures that are essential for network traffic monitoring. In this paper we have proposed a novel method for network traffic measurement and analysis.

Author: Department of Computer Science, Adamas Institute of Technology, Kolkata, India. e-mail: tanmaypaulcs@gmail.com

II. Software Overview

In Hardoop we can analyses and process large data set. It eliminated the use of expensive hardware for storing and analyzing huge data. It minimizes the cost of installing distributed parallel processing of the data by installing hardware in existing servers. By implementation of hardoop it enables to process and analyses the data more efficiently and also by reducing the cost. It also enables the organization to import and use the data one became absolute. Below in Fig 1 the flow chart of data flow of 7 layers of OSI model of traffic analysis based on hardoop is given. The tools mentioned below in Table 1 are efficient of data analyzing but are limited to storage and measurement. The traffic sampling method can be used to overcome the limitation where results are drawn through partial observation. The implementation of SQL is also not proposed due to its nature of query operation. Below in Table 1 networking traffic monitoring tools are given with uses and limitations are described.

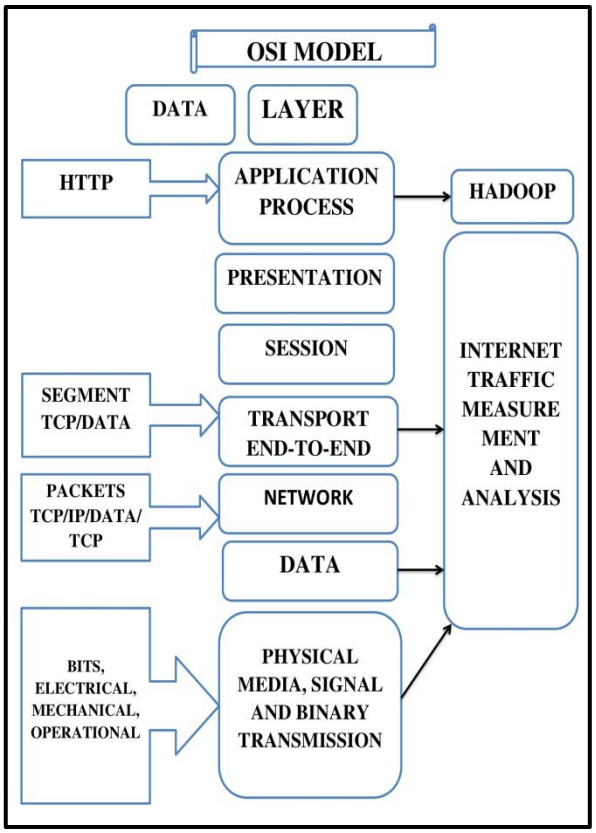


Fig.1: OSI model based on hardoop (source: 6)

Table 1: Networking traffic monitoring tools.

TOOL NAME	OPERATING SYSTEM	LANGUAGE	USE	DISADVANTAGE
NETWORK MINER	Windows, Mac, Linux, FreeBSD.	C	Used as passive network sniffer/packet capturing tool in order to detect OS sessions, host name, open ports etc.	Cost is high about 70$
WIRESHARK	Linux, OS X, BSD, Solaris, windows.	C, C++	It allows examining of the data from a live network or from a capture file on disk. The data can also be browsed and delving down into packet level as required	The main issue is the security features of this tool.
TEPDUMP	Unix like OS, Linux, OS X, BSD, Solaris, windows, android and AIX	C	The user with the necessary privilege acting on a router or gateway through which unencrypted traffic such as telnet passes can use Tepdump to view login id, passwords, URLs, content of website being viewed or any other unencrypted information	It does not receives new features update and keep resolving the bugs and troubleshooting the previous networking issues.

III. System Overview

The system proposed involved firstly input conversion, secondly hardoop pre-processing and qlikView [7] analysis. At first for the packet capture jpcap and wincap [8-9] is used for capturing which is used for supporting the jdk environment and wincap supports the window environment. After capturing the packet gets converted into .text file or .csv file for training data. The dataset made gets loaded as input for category. The processed file is stored in HDFS and to represent in HIVE file externally. And at last IP analysis, port no, protocol and displayed in graphical format. Below in Fig2 the work flow diagram of the proposed system has been given.

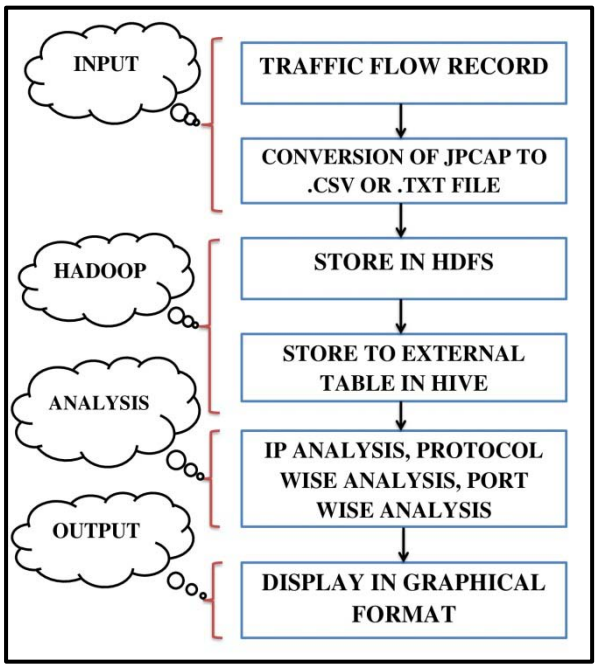


Fig.2: Work flow diagram

Hardoop has been implemented for network traffic analysis and measurement. The various characteristic of traffic data is been considered as IP address for traffic counting, total traffic data size, traffic counting with port based classification where total traffic and size per port is calculated. The internet traffic is being captured from Adamas institute, Kolkata which has been stored in jcap and wincap format. The slave node stores the data with the replication factor of 3 which means 1 file is stored and min Fig3 the network diagram has been given.

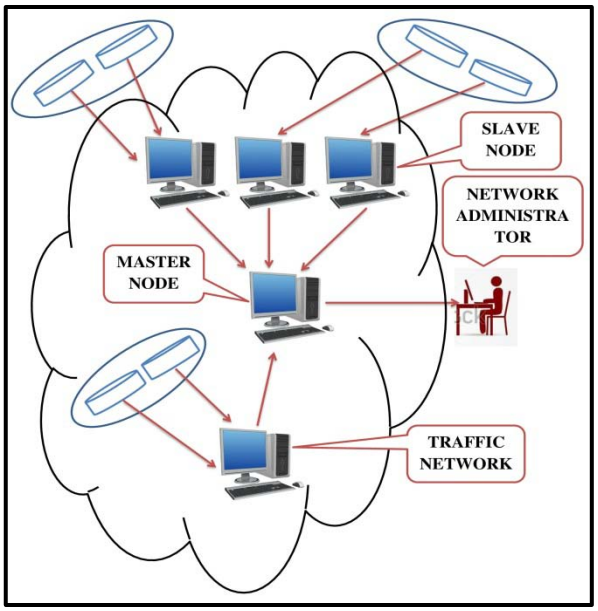


Fig.3: Network diagram.

IV. Experimental Evaluation

Protocol based network packet are captured, port number having LAN making use of java API.2 and IP addresses. The captured file stored in HDFS [10-11] is described data wise. The top 10 IP address can be calculated to define the user usage so that the network which consumes more traffic or more bandwidth can be identified. The total number of packet has also been calculated based on port which his called port-wise byte counts. Port 443 (HTTPS) having the highest number of count which is about 59% has also been shown below. The size of packet and total number of packet each day has been calculated. Below in Fig4 the top 10 IP address usage is shown and in Fig5 the port wise byte count is also shown.

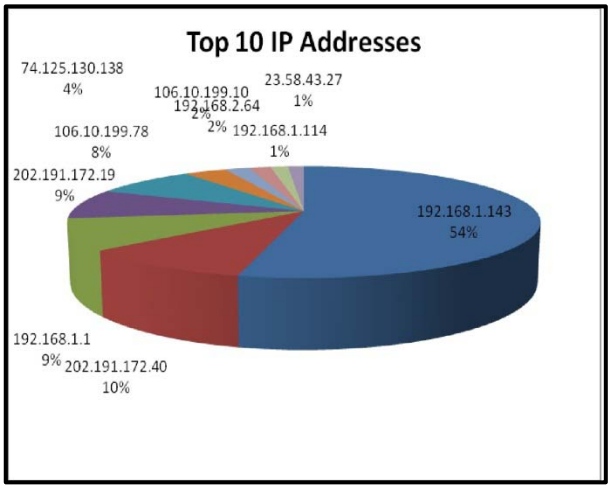


Fig.4: The top 10 IP address data usage.

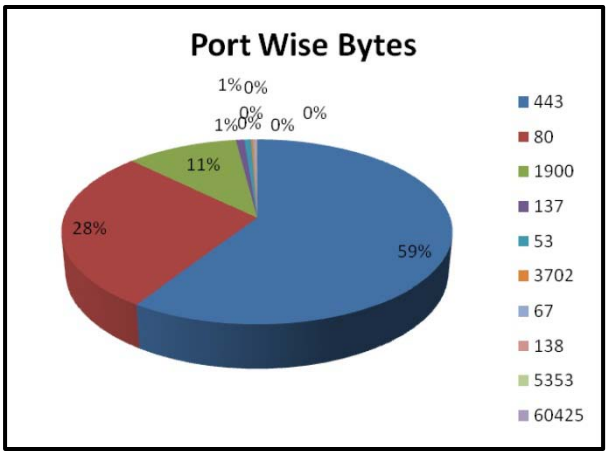


Fig.5: The port-wise byte count.

V. Conclusion

The network traffic analysis we proposed in this paper will be very efficient for the network administrator to monitor the bandwidth consumption and maintain the system and trouble shoot bugs if necessary. In the paper our main focus was on the flow packet and analysis by network topology. The huge amount of data cannot be handled with single server so large dataset is necessary for matching the computing and storage, and scalable analysis becomes a problem. That the reason we introduce Hardoop as an open-source platform which resolves all the issue in large data set analysis. We have proposed the novel method of data analysis and measurement.

References Références Referencias

1. Wu Kehe; Cheng Rui; Zhang Yingqiang; Mu Hongtao "The research on the software architecture of network packet processing based on the many-core processors":2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS).
2. Ali R. Ahmadi; Laura Kane; Robert Macdonald; GrahamAult; Pedro Almeida; Sotiris Georgiopoulos; JoseBarros; Panagiotis Papadopoulos "Active network management supporting energy storage integration into system, market and the distribution network":CIRED Workshop 2016.
3. Shuonan Shang; YongqingMeng; Jian Wang; Huixuan Li; Wei Ren; Xifan Wang; Yong Cui "Research on modeling and control strategy of modular multilevel matrix converter supplying passive networks" : 2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC).
4. Liu Mingbo; Sun Wenjie; Zhao Qianhong; Tian Zhaoping "Design and implementation of IP network traffic monitoring system": 2016 15th International Conference on Optical Communications and Networks (ICOCN).

5. Vaibhav Fanibhare; Vijay Dahake “Smart Grids Map Reduce framework using Hadoop”: 2016 3rd International Conference on Signal Processing and Integrated Networks (SPIN).
6. Lena T. Ibrahim; Rosilah Hassan; Kamsuriah Ahmad; AsrulNizamAsat “A study on improvement of internet traffic measurement and analysis usingH adoop system”: 2015 International Conference on Electrical Engineering and Informatics (ICEEI).
7. Roumianallieva; KirilAnguelov; Delyana Gashurova “Monitoring and optimization of e-Services in IT Service Desk Systems”: 2016 19th International Symposium on Electrical Apparatus and Technologies (SIELA).
8. Yong Xing Wang; Xiu Zhu Jiang; Chun Wang “Design and simulation on the PC of routing software based on Wincap”: The 2011 IEEE/ICME International Conference on Complex Medical Engineering.
9. Wenjian Xing; Yunlan Zhao; Tonglei Li “Research on the Defense Against ARP Spoofing Attacks Based on Winpcap”: 2010 Second International Workshop on Education Technology and Computer Science.
10. Mao Ye; Jun Wang; Jiangling Yin; Xuhong Zhang “Accelerating I/O Performance of SVM on HDFS”: 2016 IEEE International Conference on Cluster Computing (CLUSTER).
11. Bikash Agrawal; Raymond Hansen; Chunming Rong; Tomasz Wiktorski “SD-HDFS: Secure Deletion in Hadoop Distributed File System”: 2016 IEEE International Congress on Big Data (Big Data Congress).

Global Journal of Computer Science and Technology: C
Software & Data Engineering
Volume 17 Issue 2 Version 1.0 Year 2017
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Strategic ERP [Enterprise Resource Planning] System Planning in Alignment with Business Planning for its Improvements

By Pvl Narayana Rao & Prof (Dr.) Tapas Kumar
Lingaya's University

Abstract- Enterprise resource planning (ERP) system has been one of the greatest widespread business management systems, provided that benefits of real-time abilities and whole communication for business in large organizations. However, not all ERP systems implementations have been successful. Since ERP implementation touches entire organizations such as process, people, and culture, so on. There are a number of experiments that companies may come across in implementing ERP system. Business approach is significant to all organizations . Nearly more than 500 companies are implementing Enterprise Resource Planning (ERP) systems to improve the execution of their business strategy and to improve combination with its information technology (IT) strategy. This research observes business and IT planned alignment and pursues to determine whether an ERP implementation can drive business process reengineering and business and IT strategic alignment.

Keywords: enterprise resource planning, information technology, strategic alignment, "as-is/to-be", knowledge integration.

GJCST-C Classification: J.1

Strictly as per the compliance and regulations of:

Strategic ERP [Enterprise Resource Planning] System Planning in Alignment with Business Planning for its Improvements

Pvl Narayana Rao [α] & Prof (Dr.) Tapas Kumar [σ]

Abstract- Enterprise resource planning (ERP) system has been one of the greatest widespread business management systems, provided that benefits of real-time abilities and whole communication for business in large organizations. However, not all ERP systems implementations have been successful. Since ERP implementation touches entire organizations such as process, people, and culture, so on. There are a number of experiments that companies may come across in implementing ERP system. Business approach is significant to all organizations . Nearly more than 500 companies are implementing Enterprise Resource Planning (ERP) systems to improve the execution of their business strategy and to improve combination with its information technology (IT) strategy. This research observes business and IT planned alignment and pursues to determine whether an ERP implementation can drive business process reengineering and business and IT strategic alignment.

Keywords: enterprise resource planning, information technology, strategic alignment, "as-is/to-be", knowledge integration.

I. Introduction

a) Introduction

i. Background

K.C.P Sugar and Industries Company Ltd is unique amongst the main sugar fabricating organizations in India. Its unified business comprises of assembling and showcasing of Rectified Spirit, Extra Neutral Alcohol, Ethanol, Incidental Cogeneration of Control, Organic Dung, Mycorrhiza Vam, Calcium Lactate.

Figure 1

Author α: Research Scholar, PhD in Computer Science and Engineering, Lingaya's University, Faridabad, India. e-mail: prao9039@gmail.com
Author σ: Research Supervisor, Dean, School of Computer Science, Lingayas University, Faridabad, India. e-mail: kumartapas534@gmail.co

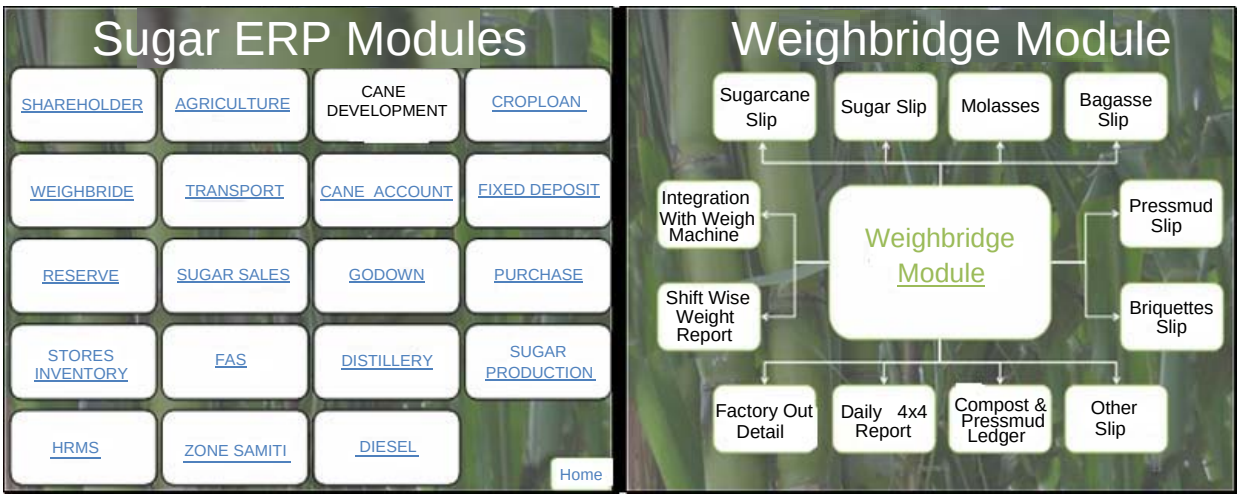


Figure 2

b) *Objectives of the Research work:* Provide a novel way to deal with SISP development definition; Establish a structure for increasing more subjective bits of knowledge into the connections of the criteria/sub criteria impacting SISP forms;.

c) *Statement of the Problem:* Since, KCP Sugar Ltd., is using ERP, It has the following significant problems were caused throughout the project in the process of abandoning its SAP software implementation. They are

1. Fake Consultants / Trainees / unqualified consulting resources on the project
2. Insufficient training and change management, poor project management and control,
3. little or no change management.
4. Waste Management Trashes Its "Fake" ERP Software

II. Literture Review

a) A Review of SISP Literature

i. *Introduction:* Goes for a decent writing audit are recognized as 'to exhibit a commonality with an assemblage of information and build up validity; to demonstrate the way of former examination and how a present venture is connected to it;

ii. *Information Feedback:* This is an outline which positions SISP in its outer surroundings (SISP is not demonstrated to; it is a subsystem of Information Systems).

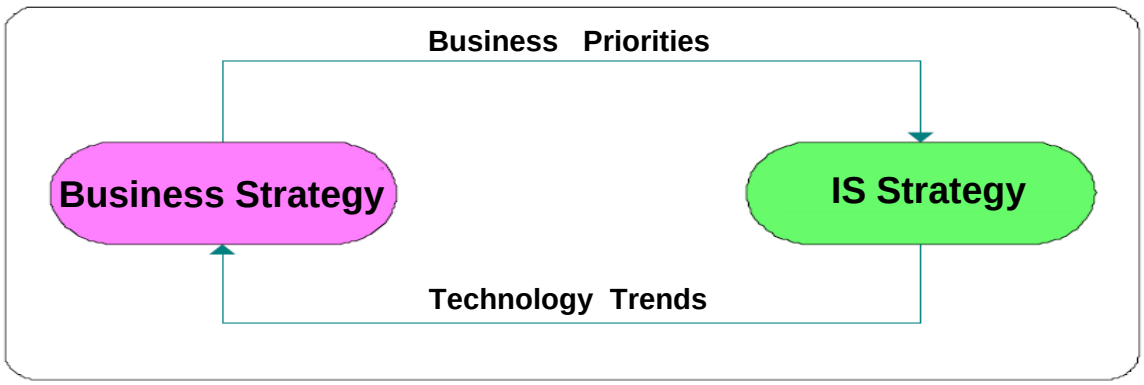


Figure 2.1: Strategic Alignment Model

III. Earlier System

a) Description of the Earlier System - Practice to be preserved from the current system

KCP Sugar and Industries Corporation Limited has detailed Code of Practices and Procedures for reasonable exposure of unpublished value touchy data and Code for Prevention of Insider Trading at the Board Meeting postponed on 29.05.2015, to be baptized as "KCP Sugar and Skills Organization Limited – Code of Fair Disclosure.

1. *Doing it in any case:* Regularly huge ERP usage ventures come up short before they even begin.
2. *No reasonable destination:* To be clear with the desires. Once an association settles on the choice to actualize another ERP framework.

3. *A decent arrangement or only an arrangement:* A point by point arrangement is exceptionally fundamental for fruitful usage. All ventures of this size begin with some sort of arrangement.
4. *Low maintenance venture administration*
5. *Under-assessing assets required:* Most basic botch to happen is with assets anticipated. Having a strong comprehension of the inward and outside assets expected to finish the undertaking is basic.

IV. Proposed System (SISP) Maturity Model

a) SISP Maturity Model

i. *Introduction:* In this Chapter, a definitional point of view of SISP development is exhibited. SISP development levels are characterized as 'Options'. At that point, taking into account the Integral Engineering approach a SISP appraisal model is produced.

ii. *UML Diagrams for Proposed SISP Maturity Model*

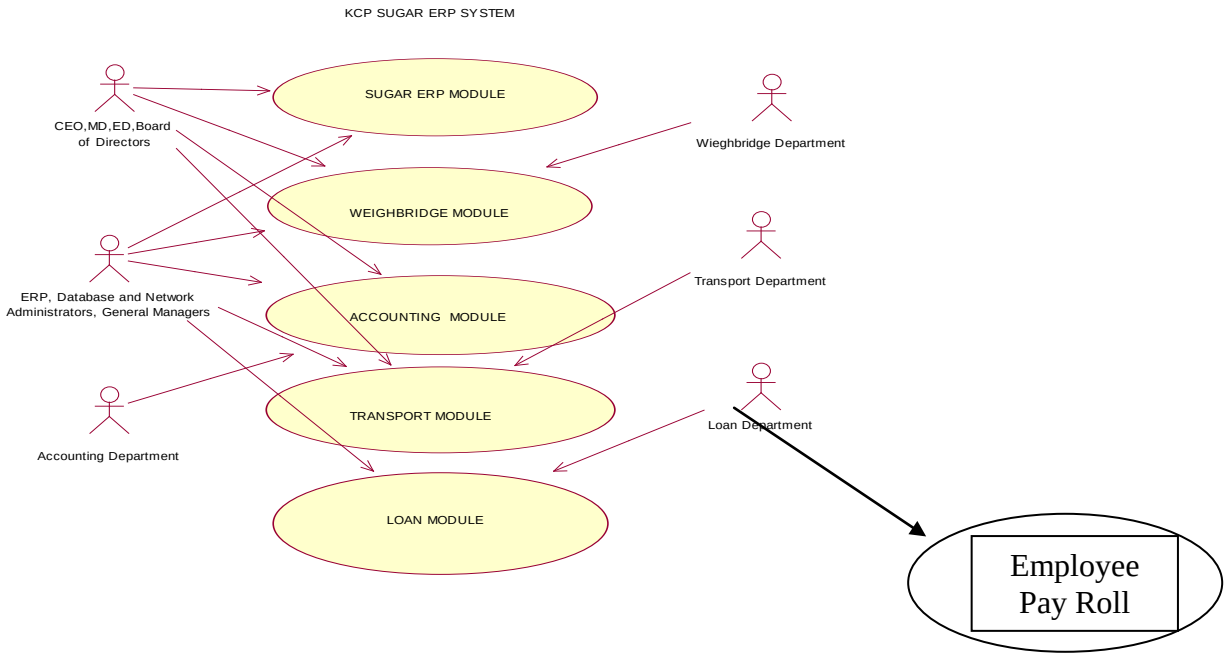


Figure 3

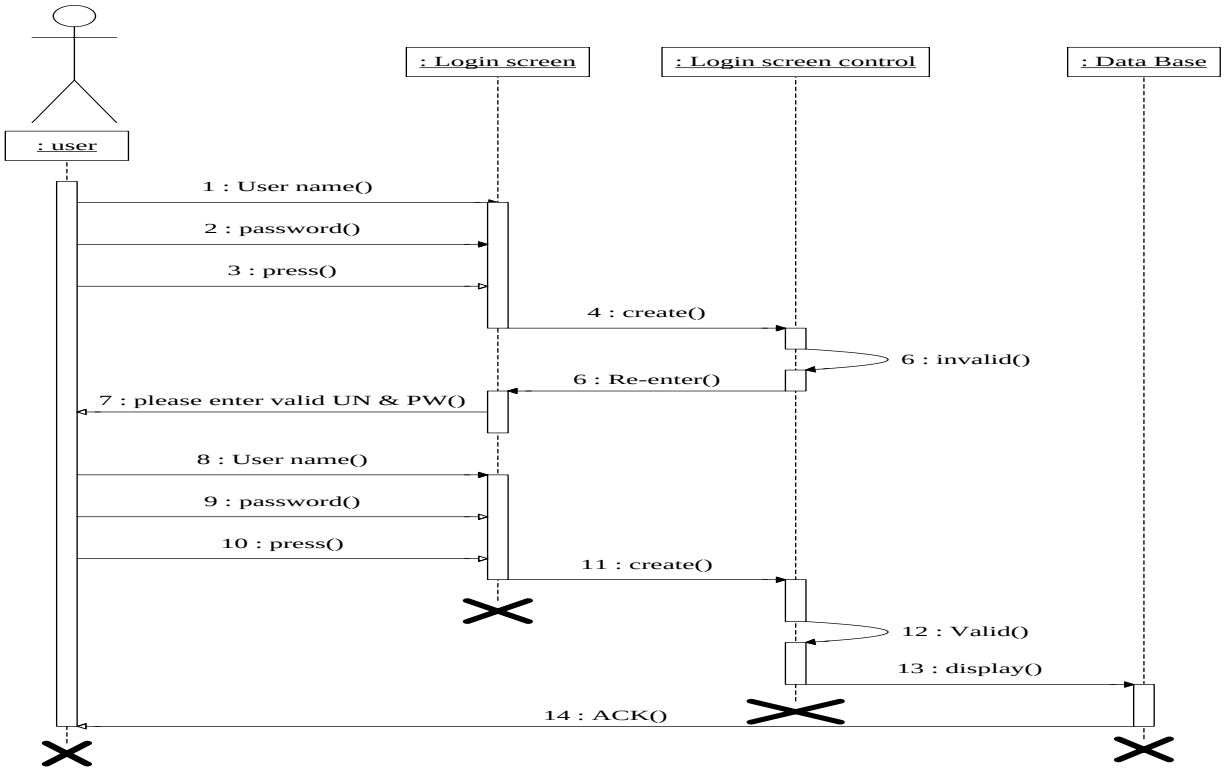


Figure 4

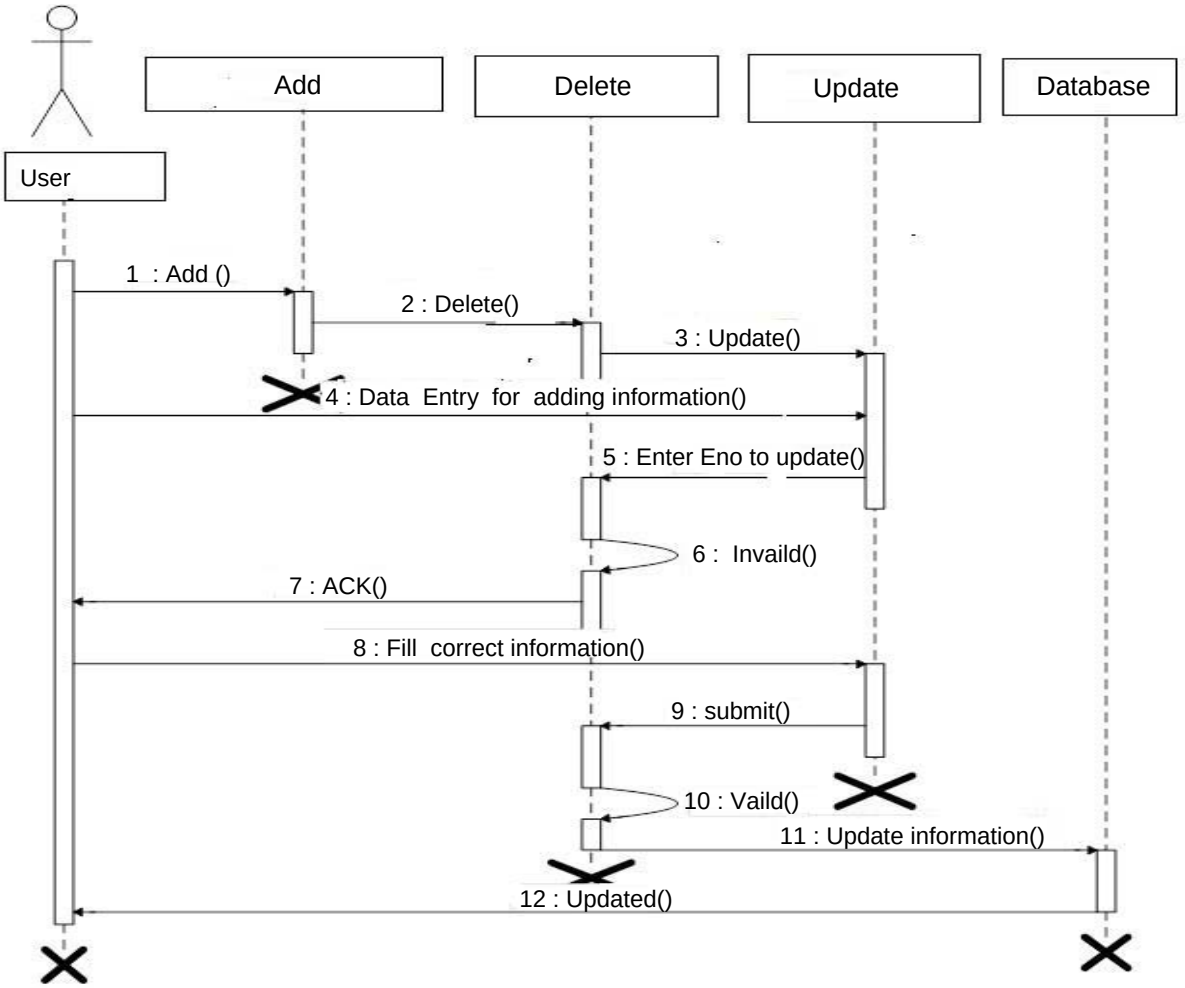


Figure 5

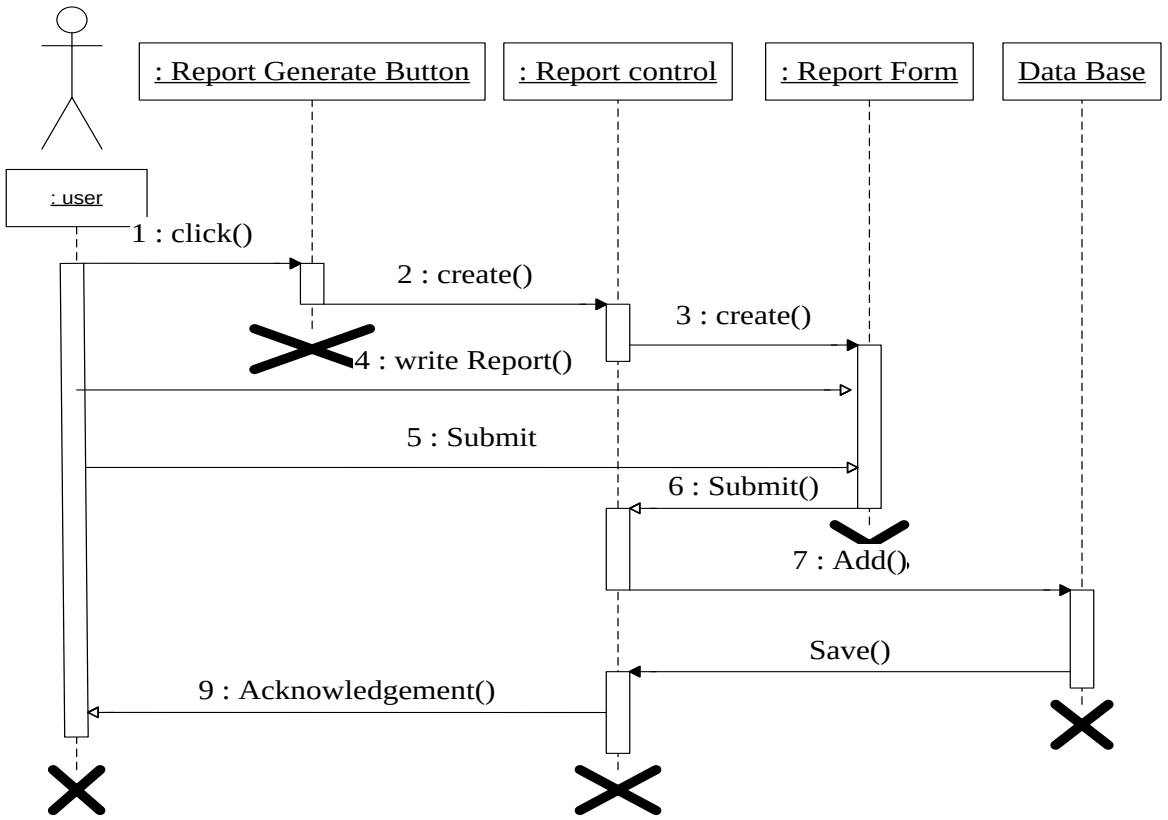


Figure 6

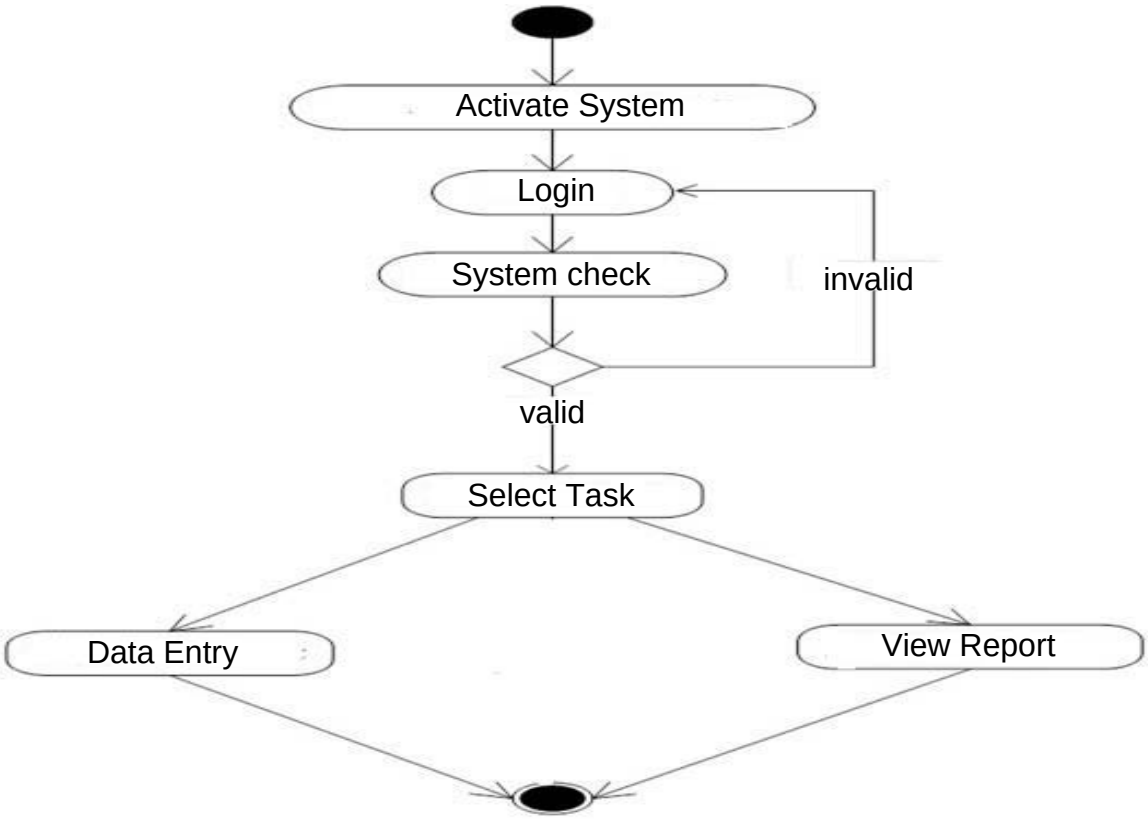


Figure 7

V. Analysis of Earlier Data and Results

a) Data Analysis

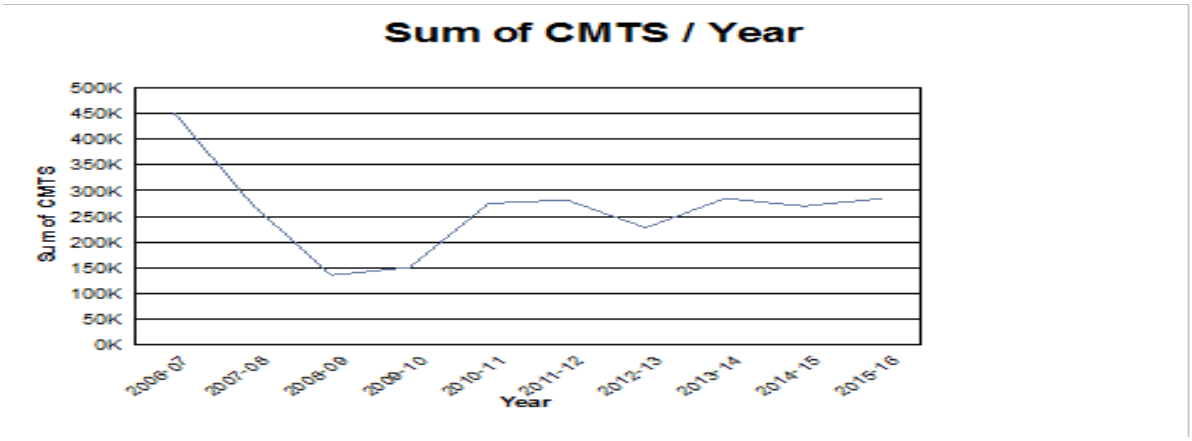
i. *Data Preparation*

Missing data and outliers are not a problem in this study; a few missing points in two questionnaires are substituted with mean values. Table 0.1 in Appendix E lists the scales against which initial items (variables) are checked for inter-correlation using principal component factor analysis (PCA).

KCP Sugar Unit, Lakshmipuram, Krishna District, Andhra Pradesh, India

SEASON WISE CANE CRUSHED, SUGAR BAGGED AND RECOVERY

SEASON	2015-16	2014-15	2013-14	2012-13	2011-12	2010-11	2009- 10	2008-09	2007-08	2006-07
Cane Crushed in MTS	2,84,686	2,70,236	2,85,464	2,27,847	2,81,847	2,75,222	1,50,759	1,35,957	2,74,193	4,53,307
Sugar bagged in QTLS	2,53,263	2,34,100	2,74,470	2,06,768	2,41,447	2,50,160	1,29,206	1,22,686	2,68,948	4,67,905
Recovery (%)	8.9	8.67	9.62	9.11	8.87	9.09	8.58	9.05	9.8	10.32



Cane Crushed in MTS

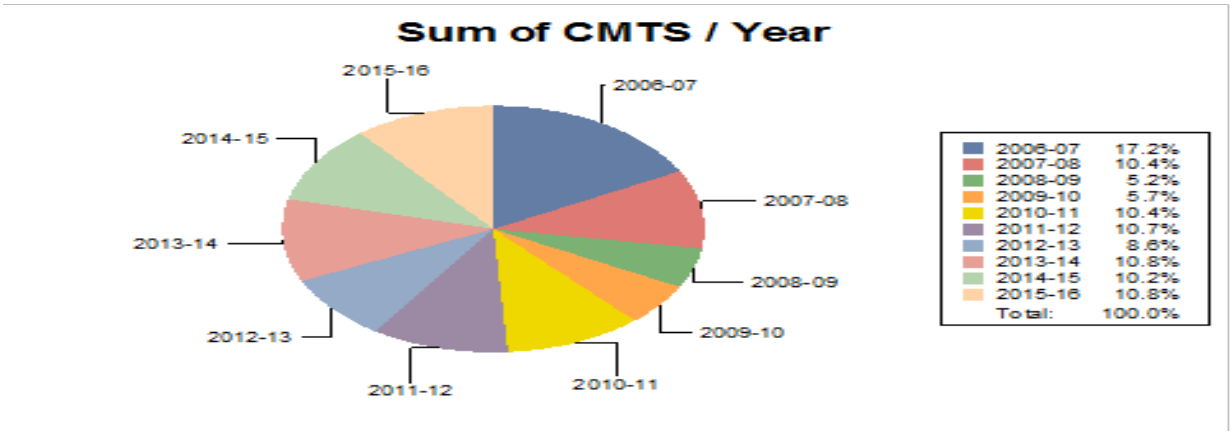


Figure 8

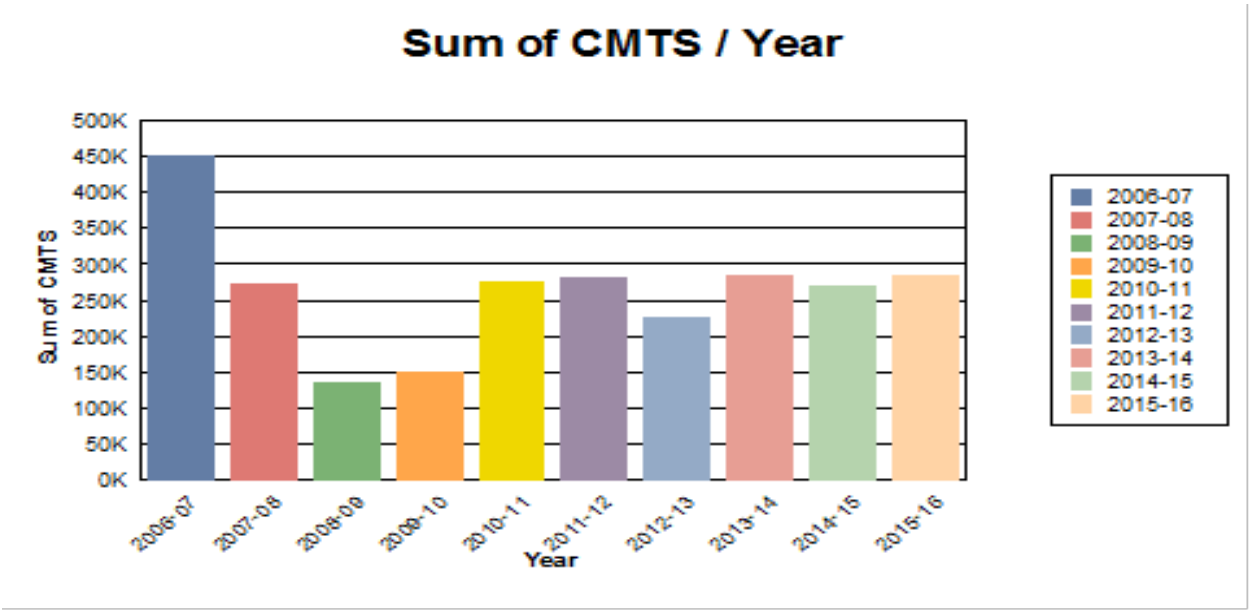


Figure 9

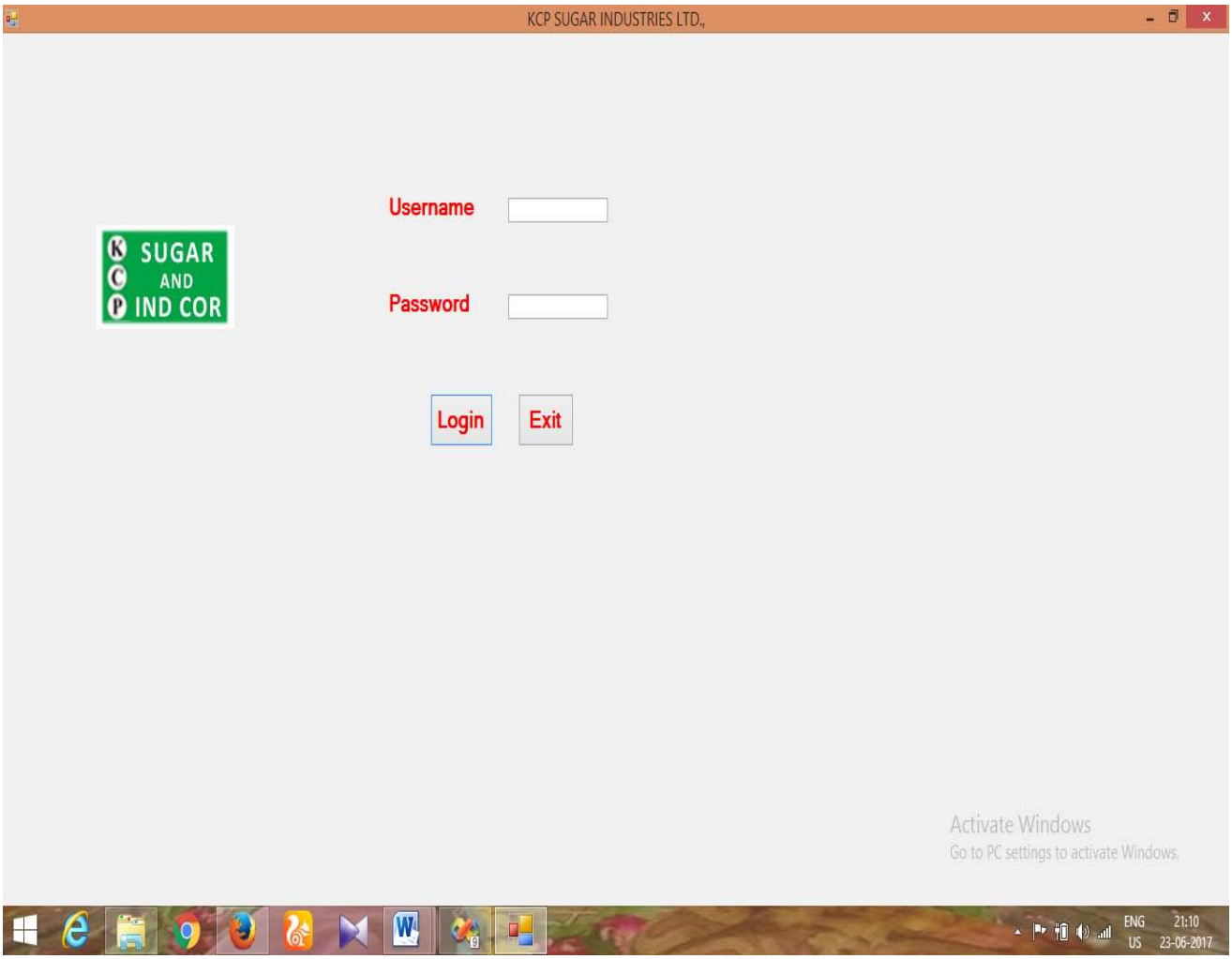


Figure 10

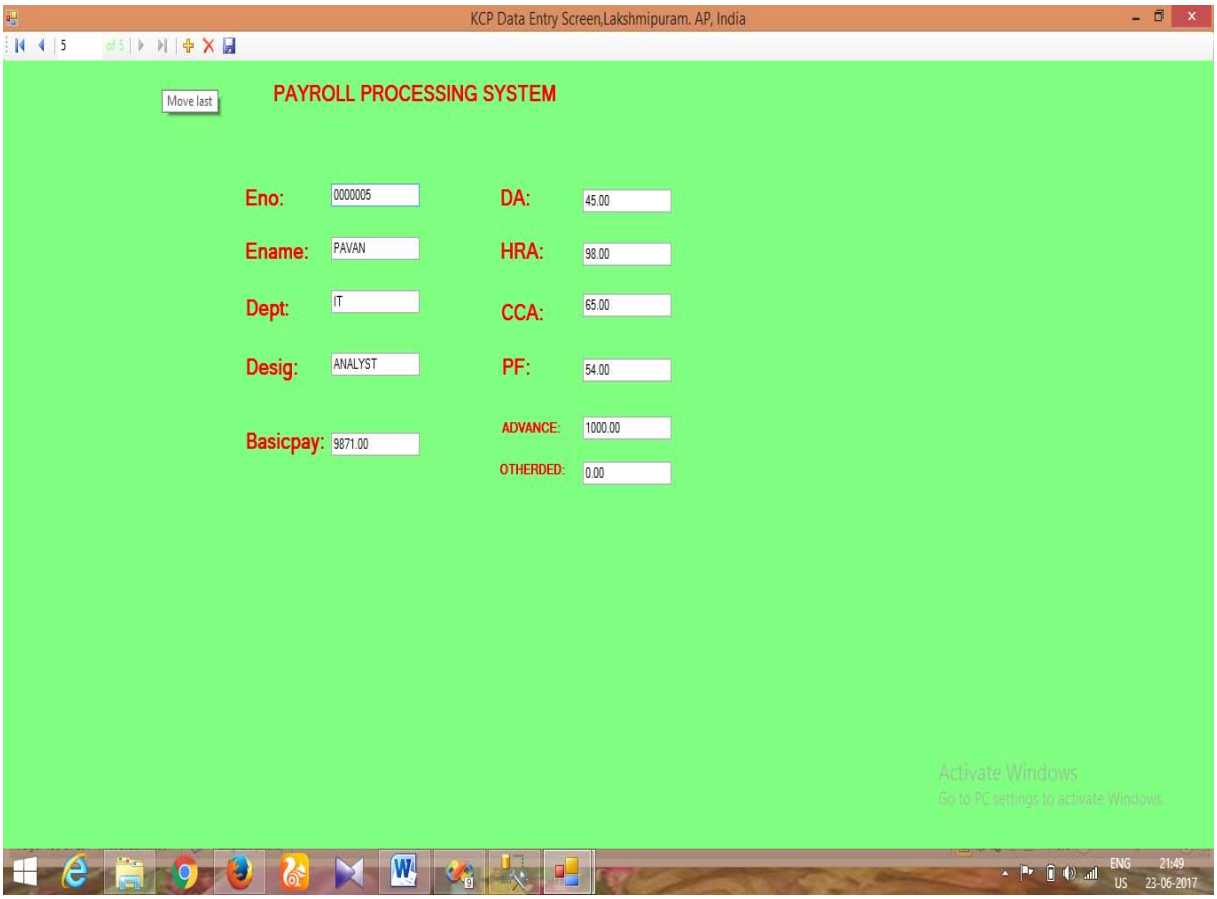


Figure 11

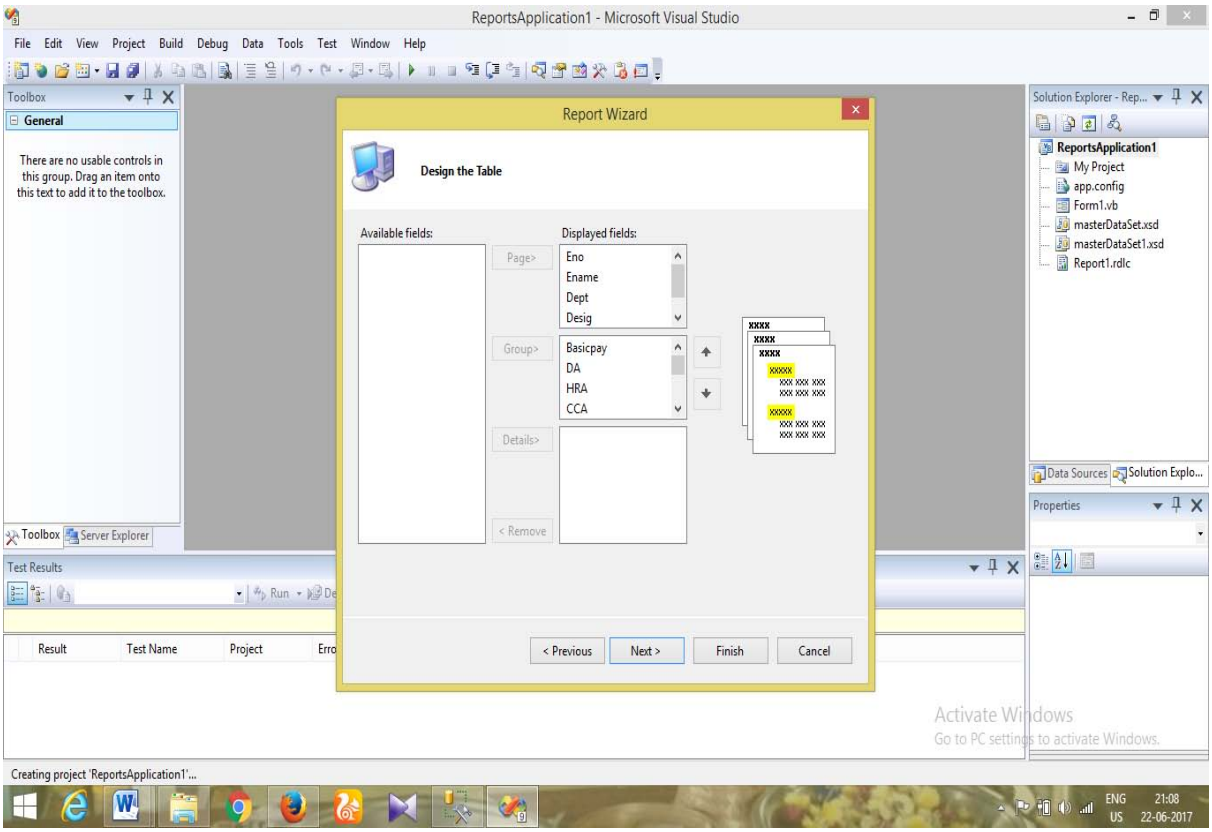


Figure 12

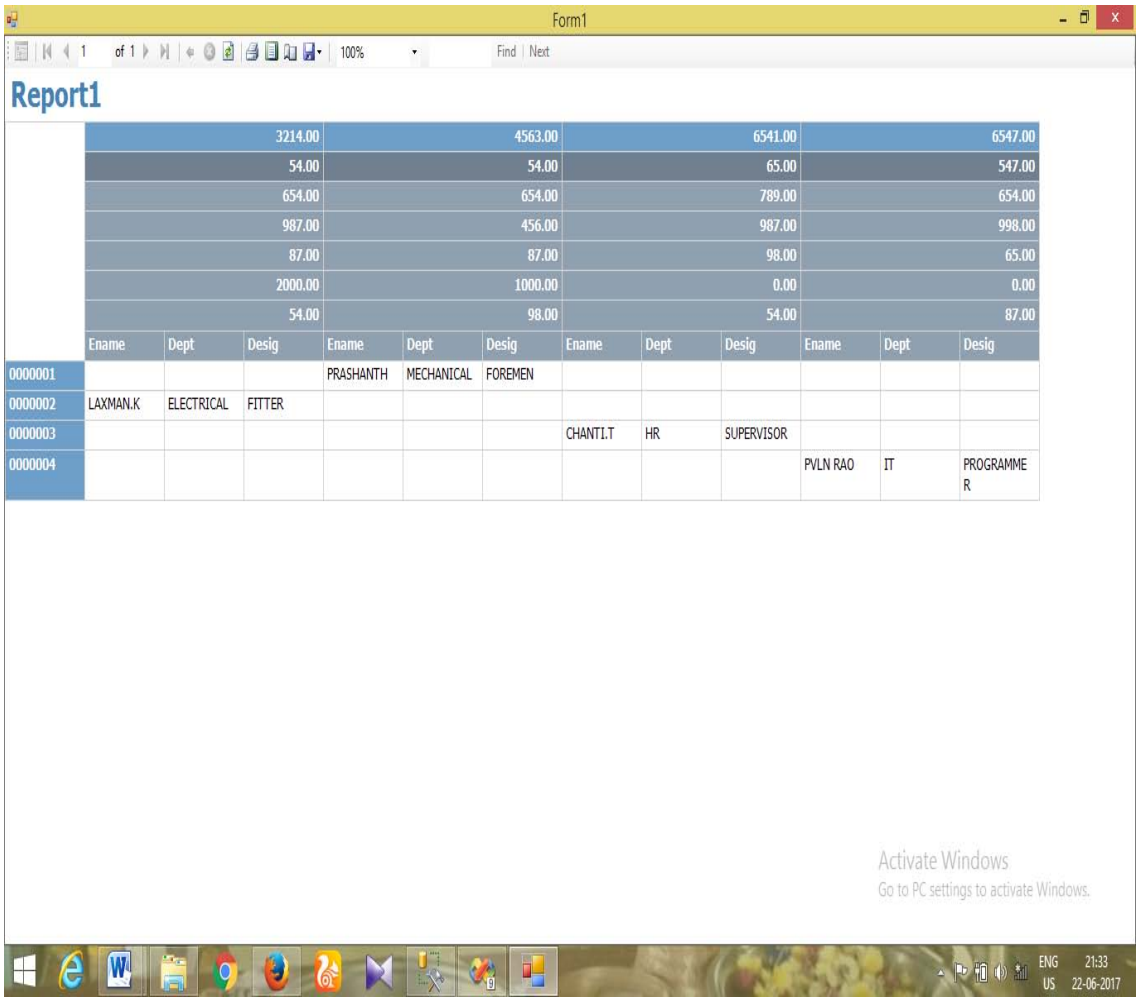


Figure 13

VI. Discussion and Conclusions

a) Summary of the Research Work

The examinations in this segment chiefly take after the characteristic method for advancement of this proposition. After a review of the primary goals and difficulties this study confronted, a rundown of the principle attributes of the models created is exhibited and after that the experimental discoveries are condensed.

b) Discussion –Revisiting the SISP Literature

The point of this segment is to analyze the experimental discoveries from Chapter 6 with the discoveries of the key references examined in Chapter 2.

c) Research Hypotheses

The model for SISP appraisal and the philosophy utilized gave a way to increasing more subjective experiences into the connections of the variables affecting the SISP process.

d) Implications for SISP Theory

One of the primary commitments to the SISP hypothesis is demonstrating the need to expand the SISP hypothesis by looking into the development of key IS arranging process essentially, i.e. isolation of SISP development from IS/IT and an association's development.

e) Limitations of the Research

All through this study particular constraints are highlighted. Here is a synopsis of confinements which naturally apply to this sort of examination.

VII. Future Expansion

a) Future Expansions

Further research in adjusting the apparatuses taking into account this structure particularly for the appraisal and estimation could augment the utilization of the instruments for proactive and responsive (feed forward and criticism) control of SISP procedures.

References Références Referencias

1. Allen, B.R. & Boyton, A.C. (1991) 'Information Architecture: In Search of Efficient Flexibility', MIS Quarterly, vol. 15, no. 4, pp. 435-445.
2. Andersen, J.T. (2001) 'Information technology, strategic decision making approach and organizational performance in different industrial settings', Journal of Strategic Information System, pp.101-119.
3. Anderson, J.C. & Gerbing, D. W. (1988)'Structural equation modeling in practice: A review and recommended two-step approach', Psychological Bulletin, vol. 103 no. 3, pp. 411-423.
4. Anderson, D.R., Sweeney, D.J. & Williams, T.A. (2004) Essentials of Modern Business Statistics. Thomson South-Western, India.
5. Andrews, D.C. & Stalick, S.K. (1994) Business reengineering–the survival guide. Prentice Hall, London.
6. Ang, J., Shaw, N. & Pavri, F. (1995) 'Identifying Strategic Management Information Systems Planning Parameters Using Case Studies', International Journal of Information Management, vol.15, no. 6, pp. 463-474.
7. Anthony, R.N. (1965) Planning and Control Systems, a Framework for Analysis. Harvard University Press, Boston, MA.
8. Applegate, L.M., Cash, J.I. & Mills, Q.D. (1988) 'Information Technology and Tomorrow's Manager', Harvard Business Review, vol. 66, no. 6.
9. Arbuckle, J.L. & Wothke, W. (1999) Amos 4.0 User Guide. Small weather Corporation, Chicago.

This page is intentionally left blank

Global Journal of Computer Science and Technology: C
Software & Data Engineering
Volume 17 Issue 2 Version 1.0 Year 2017
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

Potential of Big Data Analytics in Bio-Medical and Health Care Arena: An Exploratory Study

By Mandava Geetha Bhargava, Modugula T S Srinivasa Reddy, Shaik Shahbaz, P. Venkateswara Rao & V.Sucharita

KL University

Abstract- With the leveraging emerging big Data in every industry, Big Data can amalgamate all data related to patient to get a complete view of patient to analyze and predict the outcomes. Using big data analytics as tools. It can enhance development in new drugs, health care financing process and clinical approaches and extends a lots of benefits such as better health care quality and efficiency, fraud detection and early disease detection by means of analytics of big data. This paper provides a general survey of current progress and advances in research arena of big data, bio-medical and health care and some major challenges of big data concept and characteristics, this concerns includes big data from bio-medical and health care arena, benefits of big data, its applications and opportunities, Methods and technology progress about big data in bio-medical and health care and challenges of big data in both bio-medical and healthcare are also discussed.

Keywords: big data; health care; genomics; big data analytics.

GJCST-C Classification: J.3

Strictly as per the compliance and regulations of:

Potential of Big Data Analytics in Bio-Medical and Health Care Arena: An Exploratory Study

Mandava Geetha Bhargava [α], Modugula T S Srinivasa Reddy [σ], Shaik Shahbaz [ρ] , P. Venkateswara rao [ω] & V.Sucharita[¥]

Abstract- With the leveraging emerging big Data in every industry, Big Data can amalgamate all data related to patient to get a complete view of patient to analyze and predict the outcomes. Using big data analytics as tools. It can enhance development in new drugs, health care financing process and clinical approaches and extends a lots of benefits such as better health care quality and efficiency, fraud detection and early disease detection by means of analytics of big data. This paper provides a general survey of current progress and advances in research arena of big data, bio-medical and health care and some major challenges of big data concept and characteristics, this concerns includes big data from bio-medical and health care arena, benefits of big data, its applications and opportunities, Methods and technology progress about big data in bio-medical and health care and challenges of big data in both bio-medical and healthcare are also discussed.

Keywords: big data; health care; genomics; big data analytics.

I. Introduction

Big Data can be termed as massive data or complex that exceeds the processing capacity of traditional data processing applications and challenges are Acquire, Process, Manage, Generate, Capture, storage, sharing and visualize. Now a days Big Data is processed for analytics of various parameters in each and every field of work like Research, Education, I.T, Banking, Bio-Medical, Health Care, Construction, and Manufacturing etc. With help of some big data technologies it is been processed and characteristics of big data are 6V's i.e. Volume, Variety, Velocity, Value, Veracity, Variability as shown in fig.1.

Volume: This represents the amount of data size where it is shown through Zettabytes (Approximately 10^{15} Megabytes), Petabytes (Approximately 10^{9} Megabytes), and Terabytes (Approximately 10^{6} Megabytes), etc.

Author α σ: M.Tech Scholars of Computer Science and Engineering Department, KL University, Vaddeswaram, Guntur Dst-522502, Andhra Pradesh, India. e-mail: geetha.bhargava@hotmail.com

Author ρ: B.Tech Scholars of Computer Science and Engineering Department, KL University, Vaddeswaram, Guntur Dst-522502, Andhra Pradesh, India.

Author ω: Professor, Dept of CSE NEC, Gudur. e-mail: jesuchi78@yahoo.com

Author ¥: Professor, Dept of CSE, NEC, Gudur. e-maii: jesuchi78@yahoo.com

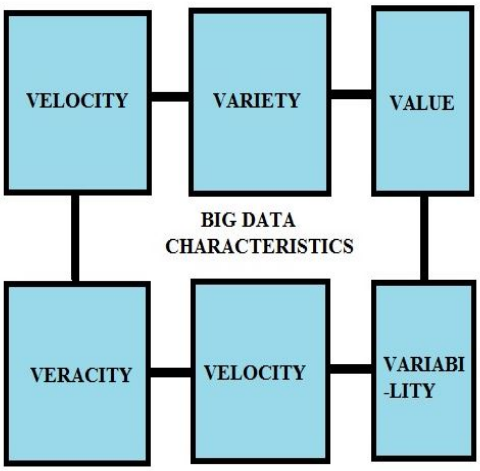


Fig.1: 6 V'S Characteristics of Big Data

Velocity: This represents the generation of data with respect to time, in other term data generated with speed.

Variety: This represents the different types of data such as unstructured data from email message, net clicks, social media streamed videos and audios and structured data from relational tables and semi-structured data from key-value web clicks, etc.

Value: Value can be defined as the collected data can bring added value. It refers to the data having value to create knowledge i.e. there is some valuable information within the collected data.

Variability: this represents the changes of data during processing and throughout lifecycle. Increase in variety and variability characterized data also increase attractiveness of data.

Veracity: this represents the two aspects of data as follows: data consistency and data trustworthiness. Sometimes the data can be in doubt due to incompleteness, ambiguities, uncertainty and deception, etc.

Big data in bio-medical and health care arena mostly denoted as electronic health data where it is large size and complex in nature to manage and process with traditional processing applications. Some data are distinguished by a means of rightness and timeliness for example: the data can generated from

wearable or implantable sensors like Bio-Metric, Blood Pressure and ECG etc. is gathered and analyzed in real time. The data in bio-medical and health care arena can be differentiate as follows: genomic data where it consists of DNA expression, genotyping and gene expression etc., clinical data where it consists of structured and unstructured data such as X-ray images, test reports, patient discharge reports and laboratory data etc., behavior and patient sentiment data consists of web data, social media data, streamed data such as telehealth and endoscopy etc., clinical reference and health publication data consists of practice guidelines, journals , medical reference material and health products etc., business, administrative and external data consists of financial data ,scheduling, billing and biometric data and other important data such as patient feedback and device data etc. as shown in figure.2.

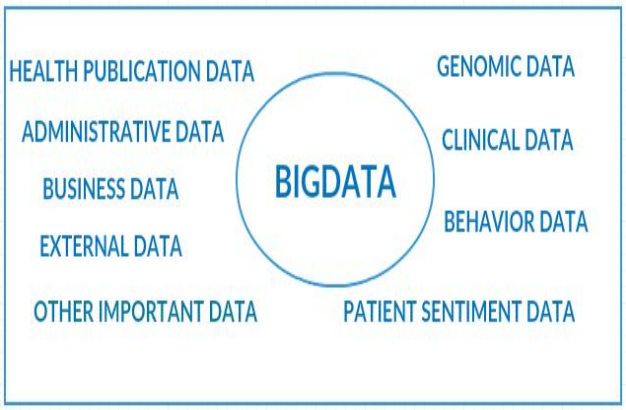


Fig.2: Different data in Bio-Medical and Health care Arena

II. Opportunities and Applications of Big Data in Biomedical and Health Care

Big data can avail support over all aspects of biomedical and health care. Big data analytics has gained traction in analytics of fraud detection and prevention, clinical outcome, genomics, epidemic disease prediction, pharmaceutical development and personalized patient care, etc. there are potential applications in biomedical and health care as they are discussed below:

a) Fraud detection and prevention analytics

Detecting, predicting and reducing fraud can be executed by using advanced analytics systems for fraud detection and checking the consistency and accuracy of claims. Big data predictive modelling can be used by health care users and payers for fraud prevention.

b) Clinical outcome Analytics

Clinical analysis can be implemented through merging financial, operational and clinical data for efficient clinical assessments. Clinical data can be utilize to reduce the manage and predict the health risks and improve clinical outcomes with cost of care.

c) Genomics analytics

The data about genes and DNAs can be analyzed for predicting and reducing the rate of disease of patients and it is becoming critical to the complete patient record to merge the both genomics data and clinical data helps to cure perilous diseases such as cancer, etc.

d) Epidemic disease prediction

In public and population health, continuously analyzing and aggregating public health data helps identifying and managing potential disease out breaks by means of analytics of social media and web-based data the disease outbreak can be known based on social content ,query activity and consumer search.

e) Design and manufacturing of medical devices

Tools of big data allows a broader set of device materials, tissue interactions, delivery methods, and anatomical configurations to be analyzed. Big data and computational methods can play an important role in design and manufacturing of medical devices.

f) Pharmaceuticals development

By analytics of pharmaceuticals data, the pharmaceutical companies can increase their ability to continue bringing new life enhancing medicines to patients in a timely manner, on basis of management of big data which was generated during all phases of pharmaceutical development, the cost of pharmaceutical product will be cost effective.

III. Research Background

Kiyana zolfaghar et al [1] done research on solutions for predicting risk of readmission for congestive heart failure patients by means integrating data of national impatient dataset and patient dataset and developed a datamining predictive model by means of integrated data and concluded that effectiveness of quality, scalability, efficiency by means of big data infrastructure on the predictive model.

Sean D.Young et al [2] done research on approaches of utilising real time social media technologies for identification and remote monitoring of HIV outcomes through negative binomial regression of tweets and concluded that the feasibility of using real time social media data to detect HIV risk-related communications, geographically map the location of those conversations and link them to national HIV outcome data for additional analysis

Priya Nambisan et al [3] done research on ruminating behavior of depression through social media, big data and public health informatics through tweets from micro blogging sites by means of screening the

vocabulary of tweets and shows sleep, pain and suicidal thoughts as they do offline and concluded that the characteristic can be used to detect and diagnose depression using the tweets in a much more effective and efficient way.

Zhendong Ji [4] done study on analysis of big data application in the medical industry and potential of its commercial value for the health care industry and concluded that by big data analysis in the medical industry provides future and promoting continuous development of medical industry through meta-analysis of gathered data.

Quan Zou et al [5] done a study on map reduce frame operation in bioinformatics through different applications and mechanisms of MapReduce and concluded that Hadoop framework has capable of handling bioinformatics data and traditional bioinformatics resources will be redesigned to support Hadoop MapReduce for high performance computing.

Liang y et al [6] studied on big data science and its applications in health care and medical research and concluded that big data offers new opportunities and promising with challenge in every field .the collaborative network, nurturing environments, team science approach with highly trained with computational skills, domain/disease expert and interdisciplinary are crucial.

Lidong Wang et al [7] done a study on big data in medical applications and health care by means of big data concept and characteristics, health care data and major issues of big data and concluded that big data can improve the research and development, translation of new therapies and has great potential to improve medicine, guide clinicians in delivering value based care.

IV. Methods and Technology Progress in Big Data

In health care /Bio-medical arena, massive amount of data about patient's medical histories, diagnosis and responses, symptomatology to treatments and therapies are gathered. Data mining techniques can implemented to derive knowledge from the gathered data in order to either examine reporting practices or to identify new patterns in infection control data. Moreover, predictive models can be utilized as detection tools can be utilize as electronic patient record gathered for every individual person of the area.

For big data healthcare/biomedical systems, the combination of Hadoop-MapReduce framework and R language is uniquely capable of storing and analyzing wide range of healthcare data types including genomic data, financial data, electronic medical records and claims data etc. the combination of two frameworks offers availability, high scalability, statistical analysis and reliability than traditional data processing systems. In addition, intelligent functional components can be built such as surveillance, detection, notification, diagnosis and recognition etc. figure 3 shows a general framework of big data analytics.

Visual analytics presents a new area of research with big data by conceptualizing the output of complex processes. The appropriate visualization solutions to the big data examples such as real time interactive visualization and metrics dash boarding [9]. Unstructured data can be converted into form of tables is to put attributes of exchangeable image file (EXIF) tags or place analyzed data where it leads to easier at visualization process. Big data can be processed through cloud technologies where it provide us operationally, insights-clinically and in research [10]. The concept of STAAS (storage as a service), is a one of the services provided by cloud computing, which provides health care center with a massive amount of storage for processing on basis of demand at low cost. [11]. Beside general infrastructure of cloud (compute, storage, virtual machine management), the following services are required to handle big data [12]. Hadoop related frameworks and tools, specialist data analytics tools, Cluster services, massively parallel processing databases, databases /servers SQL, NoSQL and security infrastructure. Organizations used various methods to de-identification of the distance data from personal identities and preserve individual's privacy. De-identification has been seen as an important security measure to be taken under the data security and accountability principle [13].

V. Proposed Approach and System

From studies of literature, the mechanisms and methodologies are basis upon Hadoop-MapReduce Framework in above literatures where it does only analysis and processing, having disadvantage of statistics analysis where it can't does and can't stores the data. Authors proposing a system which consists of RHadoop platform, it contains both R Language and Hadoop-MapReduce framework where it can process the different types of data with statistical data and streams the data after and before analysis through separate statistical package. It can visualize the processed data as output and figure 4 shows the architecture of Hadoop and data analysis tool.

RHadoop is an open source project developed by revolution analytics, provides client side integration of Hadoop and R. It allows running a MapReduce jobs within R. Needs some packages for integration of R and

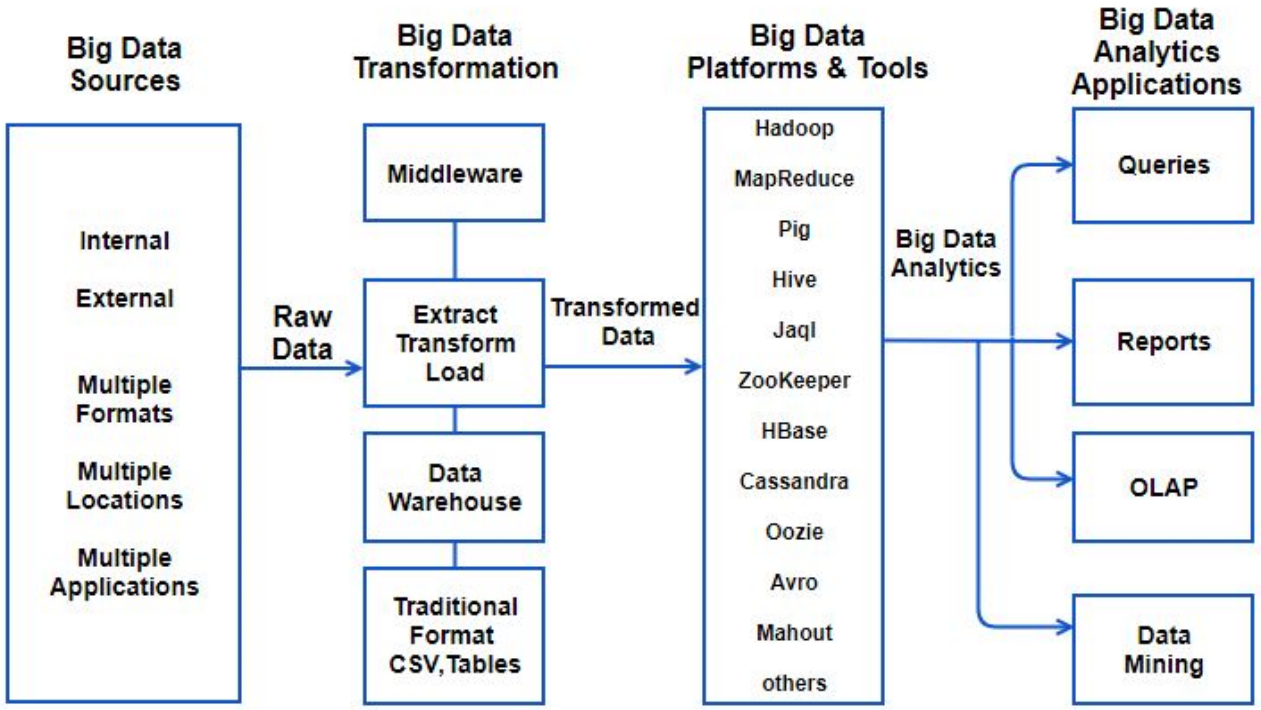


Fig.3: General Framework of Big Data Analytics [8].

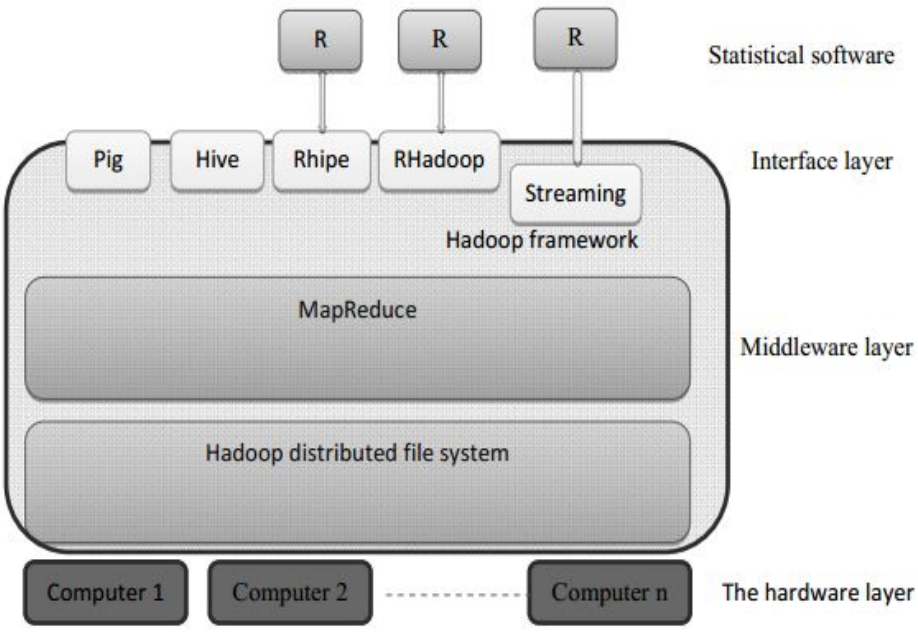


Fig.4: Architecture of Hadoop and Data Analysis

Hadoop such as plyrmr, rmr, rdfs, rhbase, etc. while setting up the RHadoop, it won't be complicated task because of dependencies on other R package.it has a wrapper R script called from streaming that invokes user defined mapper and reducer functions. Script that uses RHadoop as shown below.

```
1. library(rmr)
2. map<-function (k, v) {…}
3. reduce<-function (k,vv) {….}
4. mapreduce (input ="ffg.txt", output="output",
   textinputformat=rawtextinputformat,
   map =map, reduce=reduce)
```

firstly, the library called rmr is loaded into the memory as shown in line 1 and then it follows to definition of the map function which receives a key and value pair as input as shown in line 2. The reduce function in line 3 is called with a key and a list of values as arguments for each unique map key. Finally the script setup and runs the map reduce job as shown in line 4 in above script.

There are several layers in architecture of Hadoop and data analysis as shown in figure.4 as follows:

a) Hardware Layer
b) Middleware Layer
c) Interface Layer
d) Statistical software

a) Hardware Layer

In this layer, the components which are in this layer acts as multiple nodes on the network for analysing, processing, storing and related operations on the data by means of Hadoop and map reduce framework with Hadoop Distributed File System (HDFS) figure.5 depicts about the overview of Hadoop

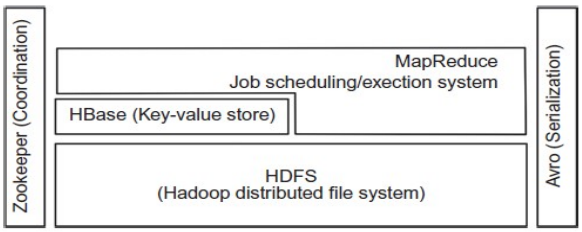


Fig.5: Overview of Hadoop

b) Middleware Layer

This layer consists of bunch of Hadoop and map reduce framework with Hadoop Distributed File System for sharing data from one node to different nodes, for processing and statistically analysis of data. Firstly the HDFS client interacts with namenode A for metadata of data stored in some datanode which are allocated to namenode A then the data can be operated by means of read or write as shown in figure.5

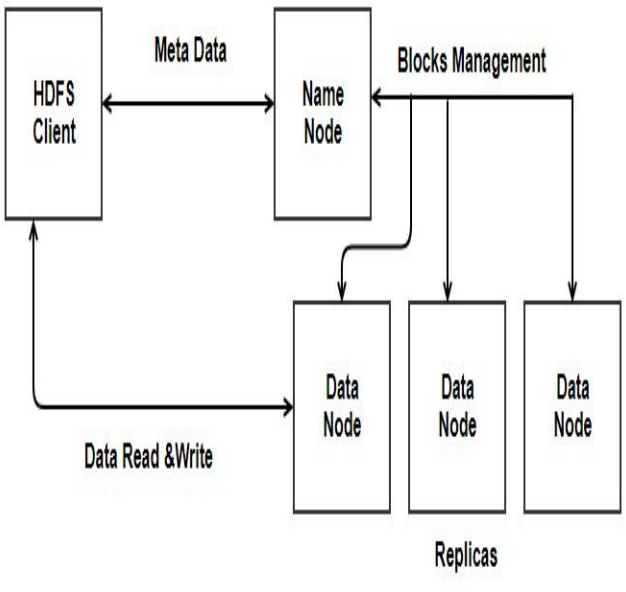


Fig.6: Architecture of Hadoop Distribution File System

c) Interface Layer

In this layer, the components consists of connectors and interfaces between Statistical analysis software and Hadoop-MapReduce framework. In this phase the data can be streamed to statistical software for analyzing and visualizing by means of direct and indirect streaming options where the direct streaming option leads to stream the data through connectors to statistical software and indirect streaming option leads to stream the data through Hadoop frameworks such as Hive, Hbase, Pig and etc. for storing and upgrading of data purpose

d) Statistical Software

The statistical software which handles data to analyze statistically through by means of visualization of graph. The software's such as SAS, R and etc.

VI. Challenges of Big Data in Biomedical and Health Care

Variety, volume and velocity characteristics of big data have brought challenges in retrieval, data storage ,visualization and search .veracity and variability of big data indicate data uncertainty and instability, which often makes big data analytics difficult and major challenges of big data in bio-medical and health care are as follows:

1) It is difficult to analyse and aggregate unstructured data such as test results, scanned documents, visual data and progress notes in patient electronic health record, etc.
2) The data in many health care providers are often segmented. Clinical data such as patient electronic

health record consists of test results, images and progress notes. Quality and outcomes data such as patient's falls, surgical site infections, etc. are in risk or quality management department where it needs standards for validating, consolidating and processing data are needed.

3) Privacy issues in the patients data such as health records , insurance details,etc. even if the privacy of the patient is protected , many health care providers are unwilling to share data due to market competition
4) Collected data can be damaged or leaked through hackers.

VII. Discussion and Conclusion

Big data is based on data generated from whole process of diagnosis and response of each case. It can lead to develop predictive models to determine which patients and health care users are mostly benefit from care management plan. By means of data analytics, it offers disease prevention, reduce in medical errors and better outcomes. it can improves and develops new therapies and research and development and has great potential to improve guide clinicians and medicines in delivering value based care. Big data has challenges in arena of bio-medical and health care to overcome such as information security, lack of infrastructure, data privacy and leakage, etc.

Utmost of all challenges can be scope for future research topics, the following topics may have a chance of future research: medical data confidentiality and interoperability, indexing and processing of continuous data, analyzing and aggregating of unstructured biomedical and health care data, security of health care data, etc. this paper focused on study on potential of big data in biomedical and health care arena and their application, challenges and opportunities.

References Références Referencias

1. Kiyana Zolfagha, Naren Meadem, Ankur Teredesai, Senjuti Basu Roy, Brian Muckian(2013), Big Data Solutions for Predicting Risk of Readmission for Congestive Heart Failure Patients, IEEE International Conference of big data,64-71
2. Sean D Young, Caitlin Rivers, Bryan Lewis (2014), methods of using real-time social media technologies for detection and remote monitoring of HIV outcomes, preventive medicine volume 63, 112-115.
3. Priya Nambisan, Zhihui Luo, Akshat Kapoor, Timothy B Patrick, Ron A Cisler, (2015) Social media, big data and public health informatics: Ruminating behaviour of depression revealed through twitter, 48th Hawaii International Conference on system sciences, 2906-2913.
4. Zhendong ji, (2015) Applications analysis of big data analysis in the medical industry, international journal of database theory and application, vol.8, No.4, 107-116.
5. Quan Zou, Xu-Bin Li, Wen-Rui Jiang, Zi-Yu Lin, Gui-Lin Li and Ke Chen (2013), Survey of Mapreduce frame operation in bioinformatics, briefings in bioinformatics, Vol.15, No.4, 637-647.
6. Liang Y and Kelemen A (2016), Big Data Science and its Applications in Healthcare and Medical Research: Challenges and Opportunities .Austin Biom and Biostat, 3 (1).
7. Lidong Wang and Cheryl Ann Alexander (2015). Big Data in Medical Applications and Health Care, American Medical Journal, 6 (1).
8. Raghupathi, W. and V. Raghupathi, (2014). Big data analytics in healthcare: Promise and potential. Health Inform. Science Syst., 2: 1-10. DOI: 10.1186/2047-2501-2-3.
9. Schultz, T., (2013). Turning healthcare challenges into big data opportunities: A use-case review across the pharmaceutical development lifecycle. Bull. Association Inform. Sci. Technol., 39: 34-40. DOI:10.1002/bult.2013.1720390508.
10. Shrestha, R.B., (2014). Big data and cloud computing. Applied Radiology.
11. Hsieh, J.C., A.H. Li and C.C. Yang, (2013). Mobile, cloud and big data computing: Contributions, challenges and new directions in tele cardiology. Int. J. Environ.Res. Public Health, 10: 6131-6153.DOI: 10.3390/I jerph10116131.
12. Turk, M., (2012). A chart of the big data ecosystem, take 2.
13. Tene, O. and J. Polonetsky, (2013). Big data for all: Privacy and user control in the age of analytics. Northwestern J. Intellectual Property, 11: 239-272.

Global Journals Inc. (US) Guidelines Handbook 2017

www.GlobalJournals.org

FELLOW OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (FARSC)

Global Journals Incorporate (USA) is accredited by Open Association of Research Society (OARS), U.S.A and in turn, awards "FARSC" title to individuals. The 'FARSC' title is accorded to a selected professional after the approval of the Editor-in-Chief/Editorial Board Members/Dean.

- The "FARSC" is a dignified title which is accorded to a person's name viz. Dr. John E. Hall, Ph.D., FARSC or William Walldroff, M.S., FARSC.

FARSC accrediting is an honor. It authenticates your research activities. After recognition as FARSC, you can add 'FARSC' title with your name as you use this recognition as additional suffix to your status. This will definitely enhance and add more value and repute to your name. You may use it on your professional Counseling Materials such as CV, Resume, and Visiting Card etc.

The following benefits can be availed by you only for next three years from the date of certification:

FARSC designated members are entitled to avail a 40% discount while publishing their research papers (of a single author) with Global Journals Incorporation (USA), if the same is accepted by Editorial Board/Peer Reviewers. If you are a main author or co-author in case of multiple authors, you will be entitled to avail discount of 10%.

Once FARSC title is accorded, the Fellow is authorized to organize a symposium/seminar/conference on behalf of Global Journal Incorporation (USA).The Fellow can also participate in conference/seminar/symposium organized by another institution as representative of Global Journal. In both the cases, it is mandatory for him to discuss with us and obtain our consent.

You may join as member of the Editorial Board of Global Journals Incorporation (USA) after successful completion of three years as Fellow and as Peer Reviewer. In addition, it is also desirable that you should organize seminar/symposium/conference at least once.

We shall provide you intimation regarding launching of e-version of journal of your stream time to time. This may be utilized in your library for the enrichment of knowledge of your students as well as it can also be helpful for the concerned faculty members.

The FARSC can go through standards of OARS. You can also play vital role if you have any suggestions so that proper amendment can take place to improve the same for the benefit of entire research community.

As FARSC, you will be given a renowned, secure and free professional email address with 100 GB of space e.g. johnhall@globaljournals.org. This will include Webmail, Spam Assassin, Email Forwarders, Auto-Responders, Email Delivery Route tracing, etc.

The FARSC will be eligible for a free application of standardization of their researches. Standardization of research will be subject to acceptability within stipulated norms as the next step after publishing in a journal. We shall depute a team of specialized research professionals who will render their services for elevating your researches to next higher level, which is worldwide open standardization.

The FARSC member can apply for grading and certification of standards of their educational and Institutional Degrees to Open Association of Research, Society U.S.A. Once you are designated as FARSC, you may send us a scanned copy of all of your credentials. OARS will verify, grade and certify them. This will be based on your academic records, quality of research papers published by you, and some more criteria. After certification of all your credentials by OARS, they will be published on your Fellow Profile link on website https://associationofresearch.org which will be helpful to upgrade the dignity.

The FARSC members can avail the benefits of free research podcasting in Global Research Radio with their research documents. After publishing the work, (including published elsewhere worldwide with proper authorization) you can upload your research paper with your recorded voice or you can utilize chargeable services of our professional RJs to record your paper in their voice on request.

The FARSC member also entitled to get the benefits of free research podcasting of their research documents through video clips. We can also streamline your conference videos and display your slides/ online slides and online research video clips at reasonable charges, on request.

The FARSC is eligible to earn from sales proceeds of his/her researches/reference/review Books or literature, while publishing with Global Journals. The FARSC can decide whether he/she would like to publish his/her research in a closed manner. In this case, whenever readers purchase that individual research paper for reading, maximum 60% of its profit earned as royalty by Global Journals, will be credited to his/her bank account. The entire entitled amount will be credited to his/her bank account exceeding limit of minimum fixed balance. There is no minimum time limit for collection. The FARSC member can decide its price and we can help in making the right decision.

The FARSC member is eligible to join as a paid peer reviewer at Global Journals Incorporation (USA) and can get remuneration of 15% of author fees, taken from the author of a respective paper. After reviewing 5 or more papers you can request to transfer the amount to your bank account.

MEMBER OF ASSOCIATION OF RESEARCH SOCIETY IN COMPUTING (MARSC)

The ' MARSC ' title is accorded to a selected professional after the approval of the Editor-in-Chief / Editorial Board Members/Dean.
The “MARSC” is a dignified ornament which is accorded to a person’s name viz. Dr. John E. Hall, Ph.D., MARSC or William Walldroff, M.S., MARSC.

MARSC accrediting is an honor. It authenticates your research activities. After becoming MARSC, you can add 'MARSC' title with your name as you use this recognition as additional suffix to your status. This will definitely enhance and add more value and repute to your name. You may use it on your professional Counseling Materials such as CV, Resume, Visiting Card and Name Plate etc.

The following benefitscan be availed by you only for next three years from the date of certification.

MARSC designated members are entitled to avail a 25% discount while publishing their research papers (of a single author) in Global Journals Inc., if the same is accepted by our Editorial Board and Peer Reviewers. If you are a main author or co-author of a group of authors, you will get discount of 10%.

As MARSC, you will be given a renowned, secure and free professional email address with 30 GB of space e.g. johnhall@globaljournals.org. This will include Webmail, Spam Assassin, Email Forwarders, Auto-Responders, Email Delivery Route tracing, etc.

We shall provide you intimation regarding launching of e-version of journal of your stream time to time. This may be utilized in your library for the enrichment of knowledge of your students as well as it can also be helpful for the concerned faculty members.

The MARSC member can apply for approval, grading and certification of standards of their educational and Institutional Degrees to Open Association of Research, Society U.S.A.

Once you are designated as MARSC, you may send us a scanned copy of all of your credentials. OARS will verify, grade and certify them. This will be based on your academic records, quality of research papers published by you, and some more criteria.

It is mandatory to read all terms and conditions carefully.

AUXILIARY MEMBERSHIPS

Institutional Fellow of Open Association of Research Society (USA)-OARS (USA)

Global Journals Incorporation (USA) is accredited by Open Association of Research Society, U.S.A (OARS) and in turn, affiliates research institutions as "Institutional Fellow of Open Association of Research Society" (IFOARS).

The "FARSC" is a dignified title which is accorded to a person's name viz. Dr. John E. Hall, Ph.D., FARSC or William Walldroff, M.S., FARSC.

The IFOARS institution is entitled to form a Board comprised of one Chairperson and three to five board members preferably from different streams. The Board will be recognized as "Institutional Board of Open Association of Research Society"-(IBOARS).

The Institute will be entitled to following benefits:

The IBOARS can initially review research papers of their institute and recommend them to publish with respective journal of Global Journals. It can also review the papers of other institutions after obtaining our consent. The second review will be done by peer reviewer of Global Journals Incorporation (USA) The Board is at liberty to appoint a peer reviewer with the approval of chairperson after consulting us.

The author fees of such paper may be waived off up to 40%.

The Global Journals Incorporation (USA) at its discretion can also refer double blind peer reviewed paper at their end to the board for the verification and to get recommendation for final stage of acceptance of publication.

The IBOARS can organize symposium/seminar/conference in their country on behalf of Global Journals Incorporation (USA)-OARS (USA). The terms and conditions can be discussed separately.

The Board can also play vital role by exploring and giving valuable suggestions regarding the Standards of "Open Association of Research Society, U.S.A (OARS)" so that proper amendment can take place for the benefit of entire research community. We shall provide details of particular standard only on receipt of request from the Board.

The board members can also join us as Individual Fellow with 40% discount on total fees applicable to Individual Fellow. They will be entitled to avail all the benefits as declared. Please visit Individual Fellow-sub menu of GlobalJournals.org to have more relevant details.

We shall provide you intimation regarding launching of e-version of journal of your stream time to time. This may be utilized in your library for the enrichment of knowledge of your students as well as it can also be helpful for the concerned faculty members.

After nomination of your institution as "Institutional Fellow" and constantly functioning successfully for one year, we can consider giving recognition to your institute to function as Regional/Zonal office on our behalf.
The board can also take up the additional allied activities for betterment after our consultation.

The following entitlements are applicable to individual Fellows:

Open Association of Research Society, U.S.A (OARS) By-laws states that an individual Fellow may use the designations as applicable, or the corresponding initials. The Credentials of individual Fellow and Associate designations signify that the individual has gained knowledge of the fundamental concepts. One is magnanimous and proficient in an expertise course covering the professional code of conduct, and follows recognized standards of practice.

Open Association of Research Society (US)/ Global Journals Incorporation (USA), as described in Corporate Statements, are educational, research publishing and professional membership organizations. Achieving our individual Fellow or Associate status is based mainly on meeting stated educational research requirements.

Disbursement of 40% Royalty earned through Global Journals : Researcher = 50%, Peer Reviewer = 37.50%, Institution = 12.50% E.g. Out of 40%, the 20% benefit should be passed on to researcher, 15 % benefit towards remuneration should be given to a reviewer and remaining 5% is to be retained by the institution.

We shall provide print version of 12 issues of any three journals [as per your requirement] out of our 38 journals worth $ 2376 USD.

Other:

The individual Fellow and Associate designations accredited by Open Association of Research Society (US) credentials signify guarantees following achievements:

- The professional accredited with Fellow honor, is entitled to various benefits viz. name, fame, honor, regular flow of income, secured bright future, social status etc.

- In addition to above, if one is single author, then entitled to 40% discount on publishing research paper and can get 10%discount if one is co-author or main author among group of authors.
- The Fellow can organize symposium/seminar/conference on behalf of Global Journals Incorporation (USA) and he/she can also attend the same organized by other institutes on behalf of Global Journals.
- The Fellow can become member of Editorial Board Member after completing 3yrs.
- The Fellow can earn 60% of sales proceeds from the sale of reference/review books/literature/publishing of research paper.
- Fellow can also join as paid peer reviewer and earn 15% remuneration of author charges and can also get an opportunity to join as member of the Editorial Board of Global Journals Incorporation (USA)
- • This individual has learned the basic methods of applying those concepts and techniques to common challenging situations. This individual has further demonstrated an in–depth understanding of the application of suitable techniques to a particular area of research practice.

Note :

"

- In future, if the board feels the necessity to change any board member, the same can be done with the consent of the chairperson along with anyone board member without our approval.
- In case, the chairperson needs to be replaced then consent of 2/3rd board members are required and they are also required to jointly pass the resolution copy of which should be sent to us. In such case, it will be compulsory to obtain our approval before replacement.
- In case of "Difference of Opinion [if any]" among the Board members, our decision will be final and binding to everyone.

"

Process of submission of Research Paper

The Area or field of specialization may or may not be of any category as mentioned in 'Scope of Journal' menu of the GlobalJournals.org website. There are 37 Research Journal categorized with Six parental Journals GJCST, GJMR, GJRE, GJMBR, GJSFR, GJHSS. For Authors should prefer the mentioned categories. There are three widely used systems UDC, DDC and LCC. The details are available as 'Knowledge Abstract' at Home page. The major advantage of this coding is that, the research work will be exposed to and shared with all over the world as we are being abstracted and indexed worldwide.

The paper should be in proper format. The format can be downloaded from first page of 'Author Guideline' Menu. The Author is expected to follow the general rules as mentioned in this menu. The paper should be written in MS-Word Format (*.DOC,*.DOCX).

The Author can submit the paper either online or offline. The authors should prefer online submission.Online Submission: There are three ways to submit your paper:

(A) (I) First, register yourself using top right corner of Home page then Login. If you are already registered, then login using your username and password.

(II) Choose corresponding Journal.

(III) Click 'Submit Manuscript'. Fill required information and Upload the paper.

(B) If you are using Internet Explorer, then Direct Submission through Homepage is also available.

(C) If these two are not convenient, and then email the paper directly to dean@globaljournals.org.

Offline Submission: Author can send the typed form of paper by Post. However, online submission should be preferred.

Preferred Author Guidelines

MANUSCRIPT STYLE INSTRUCTION (Must be strictly followed)

Page Size: 8.27" X 11'"

- Left Margin: 0.65
- Right Margin: 0.65
- Top Margin: 0.75
- Bottom Margin: 0.75
- Font type of all text should be Swis 721 Lt BT.
- Paper Title should be of Font Size 24 with one Column section.
- Author Name in Font Size of 11 with one column as of Title.
- Abstract Font size of 9 Bold, "Abstract" word in Italic Bold.
- Main Text: Font size 10 with justified two columns section
- Two Column with Equal Column with of 3.38 and Gaping of .2
- First Character must be three lines Drop capped.
- Paragraph before Spacing of 1 pt and After of 0 pt.
- Line Spacing of 1 pt
- Large Images must be in One Column
- Numbering of First Main Headings (Heading 1) must be in Roman Letters, Capital Letter, and Font Size of 10.
- Numbering of Second Main Headings (Heading 2) must be in Alphabets, Italic, and Font Size of 10.

You can use your own standard format also.
Author Guidelines:

1. General,

2. Ethical Guidelines,

3. Submission of Manuscripts,

4. Manuscript's Category,

5. Structure and Format of Manuscript,

6. After Acceptance.

1. GENERAL

Before submitting your research paper, one is advised to go through the details as mentioned in following heads. It will be beneficial, while peer reviewer justify your paper for publication.

Scope

The Global Journals Inc. (US) welcome the submission of original paper, review paper, survey article relevant to the all the streams of Philosophy and knowledge. The Global Journals Inc. (US) is parental platform for Global Journal of Computer Science and Technology, Researches in Engineering, Medical Research, Science Frontier Research, Human Social Science, Management, and Business organization. The choice of specific field can be done otherwise as following in Abstracting and Indexing Page on this Website. As the all Global

Journals Inc. (US) are being abstracted and indexed (in process) by most of the reputed organizations. Topics of only narrow interest will not be accepted unless they have wider potential or consequences.

2. ETHICAL GUIDELINES

Authors should follow the ethical guidelines as mentioned below for publication of research paper and research activities.

Papers are accepted on strict understanding that the material in whole or in part has not been, nor is being, considered for publication elsewhere. If the paper once accepted by Global Journals Inc. (US) and Editorial Board, will become the copyright of the Global Journals Inc. (US).

Authorship: The authors and coauthors should have active contribution to conception design, analysis and interpretation of findings. They should critically review the contents and drafting of the paper. All should approve the final version of the paper before submission

The Global Journals Inc. (US) follows the definition of authorship set up by the Global Academy of Research and Development. According to the Global Academy of R&D authorship, criteria must be based on:

1) Substantial contributions to conception and acquisition of data, analysis and interpretation of the findings.

2) Drafting the paper and revising it critically regarding important academic content.

3) Final approval of the version of the paper to be published.

All authors should have been credited according to their appropriate contribution in research activity and preparing paper. Contributors who do not match the criteria as authors may be mentioned under Acknowledgement.

Acknowledgements: Contributors to the research other than authors credited should be mentioned under acknowledgement. The specifications of the source of funding for the research if appropriate can be included. Suppliers of resources may be mentioned along with address.

Appeal of Decision: The Editorial Board's decision on publication of the paper is final and cannot be appealed elsewhere.

Permissions: It is the author's responsibility to have prior permission if all or parts of earlier published illustrations are used in this paper.

Please mention proper reference and appropriate acknowledgements wherever expected.

If all or parts of previously published illustrations are used, permission must be taken from the copyright holder concerned. It is the author's responsibility to take these in writing.

Approval for reproduction/modification of any information (including figures and tables) published elsewhere must be obtained by the authors/copyright holders before submission of the manuscript. Contributors (Authors) are responsible for any copyright fee involved.

3. SUBMISSION OF MANUSCRIPTS

Manuscripts should be uploaded via this online submission page. The online submission is most efficient method for submission of papers, as it enables rapid distribution of manuscripts and consequently speeds up the review procedure. It also enables authors to know the status of their own manuscripts by emailing us. Complete instructions for submitting a paper is available below.

Manuscript submission is a systematic procedure and little preparation is required beyond having all parts of your manuscript in a given format and a computer with an Internet connection and a Web browser. Full help and instructions are provided on-screen. As an author, you will be prompted for login and manuscript details as Field of Paper and then to upload your manuscript file(s) according to the instructions.

To avoid postal delays, all transaction is preferred by e-mail. A finished manuscript submission is confirmed by e-mail immediately and your paper enters the editorial process with no postal delays. When a conclusion is made about the publication of your paper by our Editorial Board, revisions can be submitted online with the same procedure, with an occasion to view and respond to all comments.

Complete support for both authors and co-author is provided.

4. MANUSCRIPT'S CATEGORY

Based on potential and nature, the manuscript can be categorized under the following heads:

Original research paper: Such papers are reports of high-level significant original research work.

Review papers: These are concise, significant but helpful and decisive topics for young researchers.

Research articles: These are handled with small investigation and applications.

Research letters: The letters are small and concise comments on previously published matters.

5. STRUCTURE AND FORMAT OF MANUSCRIPT

The recommended size of original research paper is less than seven thousand words, review papers fewer than seven thousands words also.Preparation of research paper or how to write research paper, are major hurdle, while writing manuscript. The research articles and research letters should be fewer than three thousand words, the structure original research paper; sometime review paper should be as follows:

Papers: These are reports of significant research (typically less than 7000 words equivalent, including tables, figures, references), and comprise:

(a)Title should be relevant and commensurate with the theme of the paper.

(b) A brief Summary, "Abstract" (less than 150 words) containing the major results and conclusions.

(c) Up to ten keywords, that precisely identifies the paper's subject, purpose, and focus.

(d) An Introduction, giving necessary background excluding subheadings; objectives must be clearly declared.

(e) Resources and techniques with sufficient complete experimental details (wherever possible by reference) to permit repetition; sources of information must be given and numerical methods must be specified by reference, unless non-standard.

(f) Results should be presented concisely, by well-designed tables and/or figures; the same data may not be used in both; suitable statistical data should be given. All data must be obtained with attention to numerical detail in the planning stage. As reproduced design has been recognized to be important to experiments for a considerable time, the Editor has decided that any paper that appears not to have adequate numerical treatments of the data will be returned un-refereed;

(g) Discussion should cover the implications and consequences, not just recapitulating the results; conclusions should be summarizing.

(h) Brief Acknowledgements.

(i) References in the proper form.

Authors should very cautiously consider the preparation of papers to ensure that they communicate efficiently. Papers are much more likely to be accepted, if they are cautiously designed and laid out, contain few or no errors, are summarizing, and be conventional to the approach and instructions. They will in addition, be published with much less delays than those that require much technical and editorial correction.

The Editorial Board reserves the right to make literary corrections and to make suggestions to improve briefness.

It is vital, that authors take care in submitting a manuscript that is written in simple language and adheres to published guidelines.

Format

Language: The language of publication is UK English. Authors, for whom English is a second language, must have their manuscript efficiently edited by an English-speaking person before submission to make sure that, the English is of high excellence. It is preferable, that manuscripts should be professionally edited.

Standard Usage, Abbreviations, and Units: Spelling and hyphenation should be conventional to The Concise Oxford English Dictionary. Statistics and measurements should at all times be given in figures, e.g. 16 min, except for when the number begins a sentence. When the number does not refer to a unit of measurement it should be spelt in full unless, it is 160 or greater.

Abbreviations supposed to be used carefully. The abbreviated name or expression is supposed to be cited in full at first usage, followed by the conventional abbreviation in parentheses.

Metric SI units are supposed to generally be used excluding where they conflict with current practice or are confusing. For illustration, 1.4 l rather than 1.4 × 10-3 m3, or 4 mm somewhat than 4 × 10-3 m. Chemical formula and solutions must identify the form used, e.g. anhydrous or hydrated, and the concentration must be in clearly defined units. Common species names should be followed by underlines at the first mention. For following use the generic name should be constricted to a single letter, if it is clear.

Structure

All manuscripts submitted to Global Journals Inc. (US), ought to include:

Title: The title page must carry an instructive title that reflects the content, a running title (less than 45 characters together with spaces), names of the authors and co-authors, and the place(s) wherever the work was carried out. The full postal address in addition with the e-mail address of related author must be given. Up to eleven keywords or very brief phrases have to be given to help data retrieval, mining and indexing.

Abstract, used in Original Papers and Reviews:

Optimizing Abstract for Search Engines

Many researchers searching for information online will use search engines such as Google, Yahoo or similar. By optimizing your paper for search engines, you will amplify the chance of someone finding it. This in turn will make it more likely to be viewed and/or cited in a further work. Global Journals Inc. (US) have compiled these guidelines to facilitate you to maximize the web-friendliness of the most public part of your paper.

Key Words

A major linchpin in research work for the writing research paper is the keyword search, which one will employ to find both library and Internet resources.

One must be persistent and creative in using keywords. An effective keyword search requires a strategy and planning a list of possible keywords and phrases to try.

Search engines for most searches, use Boolean searching, which is somewhat different from Internet searches. The Boolean search uses "operators," words (and, or, not, and near) that enable you to expand or narrow your affords. Tips for research paper while preparing research paper are very helpful guideline of research paper.

Choice of key words is first tool of tips to write research paper. Research paper writing is an art.A few tips for deciding as strategically as possible about keyword search:

- One should start brainstorming lists of possible keywords before even begin searching. Think about the most important concepts related to research work. Ask, "What words would a source have to include to be truly valuable in research paper?" Then consider synonyms for the important words.
- It may take the discovery of only one relevant paper to let steer in the right keyword direction because in most databases, the keywords under which a research paper is abstracted are listed with the paper.
- One should avoid outdated words.

Keywords are the key that opens a door to research work sources. Keyword searching is an art in which researcher's skills are bound to improve with experience and time.

Numerical Methods: Numerical methods used should be clear and, where appropriate, supported by references.

Acknowledgements: Please make these as concise as possible.

References

References follow the Harvard scheme of referencing. References in the text should cite the authors' names followed by the time of their publication, unless there are three or more authors when simply the first author's name is quoted followed by et al. unpublished work has to only be cited where necessary, and only in the text. Copies of references in press in other journals have to be supplied with submitted typescripts. It is necessary that all citations and references be carefully checked before submission, as mistakes or omissions will cause delays.

References to information on the World Wide Web can be given, but only if the information is available without charge to readers on an official site. Wikipedia and Similar websites are not allowed where anyone can change the information. Authors will be asked to make available electronic copies of the cited information for inclusion on the Global Journals Inc. (US) homepage at the judgment of the Editorial Board.

The Editorial Board and Global Journals Inc. (US) recommend that, citation of online-published papers and other material should be done via a DOI (digital object identifier). If an author cites anything, which does not have a DOI, they run the risk of the cited material not being noticeable.

The Editorial Board and Global Journals Inc. (US) recommend the use of a tool such as Reference Manager for reference management and formatting.

Tables, Figures and Figure Legends

Tables: Tables should be few in number, cautiously designed, uncrowned, and include only essential data. Each must have an Arabic number, e.g. Table 4, a self-explanatory caption and be on a separate sheet. Vertical lines should not be used.

Figures: Figures are supposed to be submitted as separate files. Always take in a citation in the text for each figure using Arabic numbers, e.g. Fig. 4. Artwork must be submitted online in electronic form by e-mailing them.

Preparation of Electronic Figures for Publication

Even though low quality images are sufficient for review purposes, print publication requires high quality images to prevent the final product being blurred or fuzzy. Submit (or e-mail) EPS (line art) or TIFF (halftone/photographs) files only. MS PowerPoint and Word Graphics are unsuitable for printed pictures. Do not use pixel-oriented software. Scans (TIFF only) should have a resolution of at least 350 dpi (halftone) or 700 to 1100 dpi (line drawings) in relation to the imitation size. Please give the data for figures in black and white or submit a Color Work Agreement Form. EPS files must be saved with fonts embedded (and with a TIFF preview, if possible).

For scanned images, the scanning resolution (at final image size) ought to be as follows to ensure good reproduction: line art: >650 dpi; halftones (including gel photographs) : >350 dpi; figures containing both halftone and line images: >650 dpi.

Color Charges: It is the rule of the Global Journals Inc. (US) for authors to pay the full cost for the reproduction of their color artwork. Hence, please note that, if there is color artwork in your manuscript when it is accepted for publication, we would require you to complete and return a color work agreement form before your paper can be published.

Figure Legends: Self-explanatory legends of all figures should be incorporated separately under the heading 'Legends to Figures'. In the full-text online edition of the journal, figure legends may possibly be truncated in abbreviated links to the full screen version. Therefore, the first 100 characters of any legend should notify the reader, about the key aspects of the figure.

6. AFTER ACCEPTANCE

Upon approval of a paper for publication, the manuscript will be forwarded to the dean, who is responsible for the publication of the Global Journals Inc. (US).

6.1 Proof Corrections

The corresponding author will receive an e-mail alert containing a link to a website or will be attached. A working e-mail address must therefore be provided for the related author.

Acrobat Reader will be required in order to read this file. This software can be downloaded

(Free of charge) from the following website:

www.adobe.com/products/acrobat/readstep2.html. This will facilitate the file to be opened, read on screen, and printed out in order for any corrections to be added. Further instructions will be sent with the proof.

Proofs must be returned to the dean at dean@globaljournals.org within three days of receipt.

As changes to proofs are costly, we inquire that you only correct typesetting errors. All illustrations are retained by the publisher. Please note that the authors are responsible for all statements made in their work, including changes made by the copy editor.

6.2 Early View of Global Journals Inc. (US) (Publication Prior to Print)

The Global Journals Inc. (US) are enclosed by our publishing's Early View service. Early View articles are complete full-text articles sent in advance of their publication. Early View articles are absolute and final. They have been completely reviewed, revised and edited for publication, and the authors' final corrections have been incorporated. Because they are in final form, no changes can be made after sending them. The nature of Early View articles means that they do not yet have volume, issue or page numbers, so Early View articles cannot be cited in the conventional way.

6.3 Author Services

Online production tracking is available for your article through Author Services. Author Services enables authors to track their article - once it has been accepted - through the production process to publication online and in print. Authors can check the status of their articles online and choose to receive automated e-mails at key stages of production. The authors will receive an e-mail with a unique link that enables them to register and have their article automatically added to the system. Please ensure that a complete e-mail address is provided when submitting the manuscript.

6.4 Author Material Archive Policy

Please note that if not specifically requested, publisher will dispose off hardcopy & electronic information submitted, after the two months of publication. If you require the return of any information submitted, please inform the Editorial Board or dean as soon as possible.

6.5 Offprint and Extra Copies

A PDF offprint of the online-published article will be provided free of charge to the related author, and may be distributed according to the Publisher's terms and conditions. Additional paper offprint may be ordered by emailing us at: editor@globaljournals.org .

You must strictly follow above Author Guidelines before submitting your paper or else we will not at all be responsible for any corrections in future in any of the way.

Before start writing a good quality Computer Science Research Paper, let us first understand what is Computer Science Research Paper? So, Computer Science Research Paper is the paper which is written by professionals or scientists who are associated to Computer Science and Information Technology, or doing research study in these areas. If you are novel to this field then you can consult about this field from your supervisor or guide.

TECHNIQUES FOR WRITING A GOOD QUALITY RESEARCH PAPER:

1. Choosing the topic: In most cases, the topic is searched by the interest of author but it can be also suggested by the guides. You can have several topics and then you can judge that in which topic or subject you are finding yourself most comfortable. This can be done by asking several questions to yourself, like Will I be able to carry our search in this area? Will I find all necessary recourses to accomplish the search? Will I be able to find all information in this field area? If the answer of these types of questions will be "Yes" then you can choose that topic. In most of the cases, you may have to conduct the surveys and have to visit several places because this field is related to Computer Science and Information Technology. Also, you may have to do a lot of work to find all rise and falls regarding the various data of that subject. Sometimes, detailed information plays a vital role, instead of short information.

2. Evaluators are human: First thing to remember that evaluators are also human being. They are not only meant for rejecting a paper. They are here to evaluate your paper. So, present your Best.

3. Think Like Evaluators: If you are in a confusion or getting demotivated that your paper will be accepted by evaluators or not, then think and try to evaluate your paper like an Evaluator. Try to understand that what an evaluator wants in your research paper and automatically you will have your answer.

4. Make blueprints of paper: The outline is the plan or framework that will help you to arrange your thoughts. It will make your paper logical. But remember that all points of your outline must be related to the topic you have chosen.

5. Ask your Guides: If you are having any difficulty in your research, then do not hesitate to share your difficulty to your guide (if you have any). They will surely help you out and resolve your doubts. If you can't clarify what exactly you require for your work then ask the supervisor to help you with the alternative. He might also provide you the list of essential readings.

6. Use of computer is recommended: As you are doing research in the field of Computer Science, then this point is quite obvious.

7. Use right software: Always use good quality software packages. If you are not capable to judge good software then you can lose quality of your paper unknowingly. There are various software programs available to help you, which you can get through Internet.

8. Use the Internet for help: An excellent start for your paper can be by using the Google. It is an excellent search engine, where you can have your doubts resolved. You may also read some answers for the frequent question how to write my research paper or find model research paper. From the internet library you can download books. If you have all required books make important reading selecting and analyzing the specified information. Then put together research paper sketch out.

9. Use and get big pictures: Always use encyclopedias, Wikipedia to get pictures so that you can go into the depth.

10. Bookmarks are useful: When you read any book or magazine, you generally use bookmarks, right! It is a good habit, which helps to not to lose your continuity. You should always use bookmarks while searching on Internet also, which will make your search easier.

11. Revise what you wrote: When you write anything, always read it, summarize it and then finalize it.

12. Make all efforts: Make all efforts to mention what you are going to write in your paper. That means always have a good start. Try to mention everything in introduction, that what is the need of a particular research paper. Polish your work by good skill of writing and always give an evaluator, what he wants.

13. Have backups: When you are going to do any important thing like making research paper, you should always have backup copies of it either in your computer or in paper. This will help you to not to lose any of your important.

14. Produce good diagrams of your own: Always try to include good charts or diagrams in your paper to improve quality. Using several and unnecessary diagrams will degrade the quality of your paper by creating "hotchpotch." So always, try to make and include those diagrams, which are made by your own to improve readability and understandability of your paper.

15. Use of direct quotes: When you do research relevant to literature, history or current affairs then use of quotes become essential but if study is relevant to science then use of quotes is not preferable.

16. Use proper verb tense: Use proper verb tenses in your paper. Use past tense, to present those events that happened. Use present tense to indicate events that are going on. Use future tense to indicate future happening events. Use of improper and wrong tenses will confuse the evaluator. Avoid the sentences that are incomplete.

17. Never use online paper: If you are getting any paper on Internet, then never use it as your research paper because it might be possible that evaluator has already seen it or maybe it is outdated version.

18. Pick a good study spot: To do your research studies always try to pick a spot, which is quiet. Every spot is not for studies. Spot that suits you choose it and proceed further.

19. Know what you know: Always try to know, what you know by making objectives. Else, you will be confused and cannot achieve your target.

20. Use good quality grammar: Always use a good quality grammar and use words that will throw positive impact on evaluator. Use of good quality grammar does not mean to use tough words, that for each word the evaluator has to go through dictionary. Do not start sentence with a conjunction. Do not fragment sentences. Eliminate one-word sentences. Ignore passive voice. Do not ever use a big word when a diminutive one would suffice. Verbs have to be in agreement with their subjects. Prepositions are not expressions to finish sentences with. It is incorrect to ever divide an infinitive. Avoid clichés like the disease. Also, always shun irritating alliteration. Use language that is simple and straight forward. put together a neat summary.

21. Arrangement of information: Each section of the main body should start with an opening sentence and there should be a changeover at the end of the section. Give only valid and powerful arguments to your topic. You may also maintain your arguments with records.

22. Never start in last minute: Always start at right time and give enough time to research work. Leaving everything to the last minute will degrade your paper and spoil your work.

23. Multitasking in research is not good: Doing several things at the same time proves bad habit in case of research activity. Research is an area, where everything has a particular time slot. Divide your research work in parts and do particular part in particular time slot.

24. Never copy others' work: Never copy others' work and give it your name because if evaluator has seen it anywhere you will be in trouble.

25. Take proper rest and food: No matter how many hours you spend for your research activity, if you are not taking care of your health then all your efforts will be in vain. For a quality research, study is must, and this can be done by taking proper rest and food.

26. Go for seminars: Attend seminars if the topic is relevant to your research area. Utilize all your resources.

27. Refresh your mind after intervals: Try to give rest to your mind by listening to soft music or by sleeping in intervals. This will also improve your memory.

28. Make colleagues: Always try to make colleagues. No matter how sharper or intelligent you are, if you make colleagues you can have several ideas, which will be helpful for your research.

29. Think technically: Always think technically. If anything happens, then search its reasons, its benefits, and demerits.

30. Think and then print: When you will go to print your paper, notice that tables are not be split, headings are not detached from their descriptions, and page sequence is maintained.

31. Adding unnecessary information: Do not add unnecessary information, like, I have used MS Excel to draw graph. Do not add irrelevant and inappropriate material. These all will create superfluous. Foreign terminology and phrases are not apropos. One should NEVER take a broad view. Analogy in script is like feathers on a snake. Not at all use a large word when a very small one would be sufficient. Use words properly, regardless of how others use them. Remove quotations. Puns are for kids, not grunt readers. Amplification is a billion times of inferior quality than sarcasm.

32. Never oversimplify everything: To add material in your research paper, never go for oversimplification. This will definitely irritate the evaluator. Be more or less specific. Also too, by no means, ever use rhythmic redundancies. Contractions aren't essential and shouldn't be there used. Comparisons are as terrible as clichés. Give up ampersands and abbreviations, and so on. Remove commas, that are, not necessary. Parenthetical words however should be together with this in commas. Understatement is all the time the complete best way to put onward earth-shaking thoughts. Give a detailed literary review.

33. Report concluded results: Use concluded results. From raw data, filter the results and then conclude your studies based on measurements and observations taken. Significant figures and appropriate number of decimal places should be used. Parenthetical remarks are prohibitive. Proofread carefully at final stage. In the end give outline to your arguments. Spot out perspectives of further study of this subject. Justify your conclusion by at the bottom of them with sufficient justifications and examples.

34. After conclusion: Once you have concluded your research, the next most important step is to present your findings. Presentation is extremely important as it is the definite medium though which your research is going to be in print to the rest of the crowd. Care should be taken to categorize your thoughts well and present them in a logical and neat manner. A good quality research paper format is essential because it serves to highlight your research paper and bring to light all necessary aspects in your research.

Informal Guidelines of Research Paper Writing

Key points to remember:

- Submit all work in its final form.
- Write your paper in the form, which is presented in the guidelines using the template.
- Please note the criterion for grading the final paper by peer-reviewers.

Final Points:

A purpose of organizing a research paper is to let people to interpret your effort selectively. The journal requires the following sections, submitted in the order listed, each section to start on a new page.

The introduction will be compiled from reference matter and will reflect the design processes or outline of basis that direct you to make study. As you will carry out the process of study, the method and process section will be constructed as like that. The result segment will show related statistics in nearly sequential order and will direct the reviewers next to the similar intellectual paths throughout the data that you took to carry out your study. The discussion section will provide understanding of the data and projections as to the implication of the results. The use of good quality references all through the paper will give the effort trustworthiness by representing an alertness of prior workings.

Writing a research paper is not an easy job no matter how trouble-free the actual research or concept. Practice, excellent preparation, and controlled record keeping are the only means to make straightforward the progression.

General style:

Specific editorial column necessities for compliance of a manuscript will always take over from directions in these general guidelines.

To make a paper clear

· Adhere to recommended page limits

Mistakes to evade

- Insertion a title at the foot of a page with the subsequent text on the next page
- Separating a table/chart or figure - impound each figure/table to a single page
- Submitting a manuscript with pages out of sequence

In every sections of your document

· Use standard writing style including articles ("a", "the," etc.)

· Keep on paying attention on the research topic of the paper

· Use paragraphs to split each significant point (excluding for the abstract)

· Align the primary line of each section

· Present your points in sound order

· Use present tense to report well accepted

· Use past tense to describe specific results

· Shun familiar wording, don't address the reviewer directly, and don't use slang, slang language, or superlatives

· Shun use of extra pictures - include only those figures essential to presenting results

Title Page:

Choose a revealing title. It should be short. It should not have non-standard acronyms or abbreviations. It should not exceed two printed lines. It should include the name(s) and address (es) of all authors.

Abstract:

The summary should be two hundred words or less. It should briefly and clearly explain the key findings reported in the manuscript--must have precise statistics. It should not have abnormal acronyms or abbreviations. It should be logical in itself. Shun citing references at this point.

An abstract is a brief distinct paragraph summary of finished work or work in development. In a minute or less a reviewer can be taught the foundation behind the study, common approach to the problem, relevant results, and significant conclusions or new questions.

Write your summary when your paper is completed because how can you write the summary of anything which is not yet written? Wealth of terminology is very essential in abstract. Yet, use comprehensive sentences and do not let go readability for briefness. You can maintain it succinct by phrasing sentences so that they provide more than lone rationale. The author can at this moment go straight to shortening the outcome. Sum up the study, with the subsequent elements in any summary. Try to maintain the initial two items to no more than one ruling each.

- Reason of the study - theory, overall issue, purpose
- Fundamental goal
- To the point depiction of the research
- Consequences, including definite statistics - if the consequences are quantitative in nature, account quantitative data; results of any numerical analysis should be reported
- Significant conclusions or questions that track from the research(es)

Approach:

- Single section, and succinct
- As a outline of job done, it is always written in past tense
- A conceptual should situate on its own, and not submit to any other part of the paper such as a form or table
- Center on shortening results - bound background information to a verdict or two, if completely necessary
- What you account in an conceptual must be regular with what you reported in the manuscript
- Exact spelling, clearness of sentences and phrases, and appropriate reporting of quantities (proper units, important statistics) are just as significant in an abstract as they are anywhere else

Introduction:

The **Introduction** should "introduce" the manuscript. The reviewer should be presented with sufficient background information to be capable to comprehend and calculate the purpose of your study without having to submit to other works. The basis for the study should be offered. Give most important references but shun difficult to make a comprehensive appraisal of the topic. In the introduction, describe the problem visibly. If the problem is not acknowledged in a logical, reasonable way, the reviewer will have no attention in your result. Speak in common terms about techniques used to explain the problem, if needed, but do not present any particulars about the protocols here. Following approach can create a valuable beginning:

- Explain the value (significance) of the study
- Shield the model - why did you employ this particular system or method? What is its compensation? You strength remark on its appropriateness from a abstract point of vision as well as point out sensible reasons for using it.
- Present a justification. Status your particular theory (es) or aim(s), and describe the logic that led you to choose them.
- Very for a short time explain the tentative propose and how it skilled the declared objectives.

Approach:

- Use past tense except for when referring to recognized facts. After all, the manuscript will be submitted after the entire job is done.
- Sort out your thoughts; manufacture one key point with every section. If you make the four points listed above, you will need a least of four paragraphs.

- Present surroundings information only as desirable in order hold up a situation. The reviewer does not desire to read the whole thing you know about a topic.
- Shape the theory/purpose specifically - do not take a broad view.
- As always, give awareness to spelling, simplicity and correctness of sentences and phrases.

Procedures (Methods and Materials):

This part is supposed to be the easiest to carve if you have good skills. A sound written Procedures segment allows a capable scientist to replacement your results. Present precise information about your supplies. The suppliers and clarity of reagents can be helpful bits of information. Present methods in sequential order but linked methodologies can be grouped as a segment. Be concise when relating the protocols. Attempt for the least amount of information that would permit another capable scientist to spare your outcome but be cautious that vital information is integrated. The use of subheadings is suggested and ought to be synchronized with the results section. When a technique is used that has been well described in another object, mention the specific item describing a way but draw the basic principle while stating the situation. The purpose is to text all particular resources and broad procedures, so that another person may use some or all of the methods in one more study or referee the scientific value of your work. It is not to be a step by step report of the whole thing you did, nor is a methods section a set of orders.

Materials:

- Explain materials individually only if the study is so complex that it saves liberty this way.
- Embrace particular materials, and any tools or provisions that are not frequently found in laboratories.
- Do not take in frequently found.
- If use of a definite type of tools.
- Materials may be reported in a part section or else they may be recognized along with your measures.

Methods:

- Report the method (not particulars of each process that engaged the same methodology)
- Describe the method entirely
- To be succinct, present methods under headings dedicated to specific dealings or groups of measures
- Simplify - details how procedures were completed not how they were exclusively performed on a particular day.
- If well known procedures were used, account the procedure by name, possibly with reference, and that's all.

Approach:

- It is embarrassed or not possible to use vigorous voice when documenting methods with no using first person, which would focus the reviewer's interest on the researcher rather than the job. As a result when script up the methods most authors use third person passive voice.
- Use standard style in this and in every other part of the paper - avoid familiar lists, and use full sentences.

What to keep away from

- Resources and methods are not a set of information.
- Skip all descriptive information and surroundings - save it for the argument.
- Leave out information that is immaterial to a third party.

Results:

The principle of a results segment is to present and demonstrate your conclusion. Create this part a entirely objective details of the outcome, and save all understanding for the discussion.

The page length of this segment is set by the sum and types of data to be reported. Carry on to be to the point, by means of statistics and tables, if suitable, to present consequences most efficiently.You must obviously differentiate material that would usually be incorporated in a study editorial from any unprocessed data or additional appendix matter that would not be available. In fact, such matter should not be submitted at all except requested by the instructor.

Content

- Sum up your conclusion in text and demonstrate them, if suitable, with figures and tables.
- In manuscript, explain each of your consequences, point the reader to remarks that are most appropriate.
- Present a background, such as by describing the question that was addressed by creation an exacting study.
- Explain results of control experiments and comprise remarks that are not accessible in a prescribed figure or table, if appropriate.
- Examine your data, then prepare the analyzed (transformed) data in the form of a figure (graph), table, or in manuscript form.

What to stay away from

- Do not discuss or infer your outcome, report surroundings information, or try to explain anything.
- Not at all, take in raw data or intermediate calculations in a research manuscript.
- Do not present the similar data more than once.
- Manuscript should complement any figures or tables, not duplicate the identical information.
- Never confuse figures with tables - there is a difference.

Approach

- As forever, use past tense when you submit to your results, and put the whole thing in a reasonable order.
- Put figures and tables, appropriately numbered, in order at the end of the report
- If you desire, you may place your figures and tables properly within the text of your results part.

Figures and tables

- If you put figures and tables at the end of the details, make certain that they are visibly distinguished from any attach appendix materials, such as raw facts
- Despite of position, each figure must be numbered one after the other and complete with subtitle
- In spite of position, each table must be titled, numbered one after the other and complete with heading
- All figure and table must be adequately complete that it could situate on its own, divide from text

Discussion:

The Discussion is expected the trickiest segment to write and describe. A lot of papers submitted for journal are discarded based on problems with the Discussion. There is no head of state for how long a argument should be. Position your understanding of the outcome visibly to lead the reviewer through your conclusions, and then finish the paper with a summing up of the implication of the study. The purpose here is to offer an understanding of your results and hold up for all of your conclusions, using facts from your research and generally accepted information, if suitable. The implication of result should be visibly described. Infer your data in the conversation in suitable depth. This means that when you clarify an observable fact you must explain mechanisms that may account for the observation. If your results vary from your prospect, make clear why that may have happened. If your results agree, then explain the theory that the proof supported. It is never suitable to just state that the data approved with prospect, and let it drop at that.

- Make a decision if each premise is supported, discarded, or if you cannot make a conclusion with assurance. Do not just dismiss a study or part of a study as "uncertain."
- Research papers are not acknowledged if the work is imperfect. Draw what conclusions you can based upon the results that you have, and take care of the study as a finished work
- You may propose future guidelines, such as how the experiment might be personalized to accomplish a new idea.
- Give details all of your remarks as much as possible, focus on mechanisms.
- Make a decision if the tentative design sufficiently addressed the theory, and whether or not it was correctly restricted.
- Try to present substitute explanations if sensible alternatives be present.
- One research will not counter an overall question, so maintain the large picture in mind, where do you go next? The best studies unlock new avenues of study. What questions remain?
- Recommendations for detailed papers will offer supplementary suggestions.

Approach:

- When you refer to information, differentiate data generated by your own studies from available information
- Submit to work done by specific persons (including you) in past tense.
- Submit to generally acknowledged facts and main beliefs in present tense.

Please carefully note down following rules and regulation before submitting your Research Paper to Global Journals Inc. (US):

Segment Draft and Final Research Paper: You have to strictly follow the template of research paper. If it is not done your paper may get rejected.

- The **major constraint** is that you must independently make all content, tables, graphs, and facts that are offered in the paper. You must write each part of the paper wholly on your own. The Peer-reviewers need to identify your own perceptive of the concepts in your own terms. NEVER extract straight from any foundation, and never rephrase someone else's analysis.

- Do not give permission to anyone else to "PROOFREAD" your manuscript.

- Methods to avoid Plagiarism is applied by us on every paper, if found guilty, you will be blacklisted by all of our collaborated research groups, your institution will be informed for this and strict legal actions will be taken immediately.)
- To guard yourself and others from possible illegal use please do not permit anyone right to use to your paper and files.

Criterion for Grading a Research Paper (Compilation) by Global Journals Inc. (US)

Please note that following table is only a Grading of "Paper Compilation" and not on "Performed/Stated Research" whose grading solely depends on Individual Assigned Peer Reviewer and Editorial Board Member. These can be available only on request and after decision of Paper. This report will be the property of Global Journals Inc. (US).

Topics	Grades		
	A-B	**C-D**	**E-F**
Abstract	Clear and concise with appropriate content, Correct format. 200 words or below	Unclear summary and no specific data, Incorrect form Above 200 words	No specific data with ambiguous information Above 250 words
Introduction	Containing all background details with clear goal and appropriate details, flow specification, no grammar and spelling mistake, well organized sentence and paragraph, reference cited	Unclear and confusing data, appropriate format, grammar and spelling errors with unorganized matter	Out of place depth and content, hazy format
Methods and Procedures	Clear and to the point with well arranged paragraph, precision and accuracy of facts and figures, well organized subheads	Difficult to comprehend with embarrassed text, too much explanation but completed	Incorrect and unorganized structure with hazy meaning
Result	Well organized, Clear and specific, Correct units with precision, correct data, well structuring of paragraph, no grammar and spelling mistake	Complete and embarrassed text, difficult to comprehend	Irregular format with wrong facts and figures
Discussion	Well organized, meaningful specification, sound conclusion, logical and concise explanation, highly structured paragraph reference cited	Wordy, unclear conclusion, spurious	Conclusion is not cited, unorganized, difficult to comprehend
References	Complete and correct format, well organized	Beside the point, Incomplete	Wrong format and structuring

Index

save our planet

Global Journal of Computer Science and Technology

Visit us on the Web at www.GlobalJournals.org | www.ComputerResearch.org
or email us at helpdesk@globaljournals.org

ISSN 9754350