RESEARCH ARTICLE OPEN ACCESS

Classification of Big Data Based Social Internet of Things

Dontala Kiran Kumar [1], Suragani.Chanti [2], Kaligiti Satyanarayana Murthy [3]
M.tech (CSE) [1], Asst Professor [2] [3], Baba institute of Technology and Sciences (BITS VIZAG)
Bakkannapalem, PM palem, Madhurawada, Visakhapatnam.

ABSTRACT
Internet of Things (IoT) is blasting and has been along with us all around and the things do not just communicate with each other without human mediation yet additionally declare some relationships called Social IoT (SIoT). The merger of the IoT and social networking, the SIoT, has made this possible. In this proposed system presented the SIoT with big data map-reduce framework along with supervised classifier model. Moreover, Gabor filter is used reduce noise and unwanted data from the database, further Hadoop MapReduce to mapping and reducing big database to improve the efficiency of the proposed work. The validity of the proposed algorithm is assessed and results are compared to existing literature to prove the performance of the proposed model.

Keywords: Social Internet of thing (SIoT), Big Data, Feature Selection, Noise, and Classification

I. INTRODUCTION

The Internet of Things (IoT) has built up a lot as of late, and the Social Internet of Things (SIoT), another utilization of IoT, is currently developing. The SIoT is a larger social network, connecting people and people, people and objects, and objects and objects [1]. One of the challenges in these scenarios is to enable the connection of everyday objects to the Internet. However, the IoT is not only about connectivity, it is about the pervasive collection and sharing of data towards a common goal [2-6]. Developing technologies and solutions for enabling the IoT vision is a major challenge. In this work, are particularly interested in identifying risk issues arising when discovering and integrating data from within IoT [7-9].

The essential thought behind a SIoT organize is that every device can look for its foreseen service by utilizing its connections, discovering its companions and companions of companions [10-12]. In the IoT, data refers to attribute values such as variables and to integer values; events refer to when certain conditions are met or when certain states are reached. Services allow certain functions to be carried out through a predefined interface [12-16]. The immense of IoT-based solutions are driving an ever-increasing number of information into endeavors; hence big data examination has become a vital component of valuable information [17-20]. This represents an opportunity that also brings several challenges to data processing systems for improving data collection, cleaning and storage, and performing real-time analytics [21-26]. The potential chance in big data processing and examination in the IoT and feature the part of information investigation in IoT applications [27-30]. Moreover, in the current scenario of Big Data, various standards and platforms have been introduced by relational database vendors. These are used for data aggregation as well as data analysis [31-35]. Consequently, it won't be confounded if the "big data and SIoT is the perfect representation of social systems and the IoT to characterize human progression [36-40].

Among feature selection algorithms, the various scheme is proposed that are classified into two broad categories, i.e., filter approaches, and wrapper approaches. In filter based techniques, before classification process, the filtration process is performed since their independent nature of usage of classification algorithms [41-43]. In the linear separable cases, Support vector machine (SVM) algorithm is put forward by the optimal classification plane.

II. LITERATURE REVIEW

Social Internet of Things (SIoTs) alludes to the quickly developing system of associated protests and individuals that can gather and exchange information utilizing embedded sensors proposed by Mohammed Zaki Hasan, and Fadi Al-Turjman in 2017 [44]. The writer proposed a bio-inspired particle multi-swarm enhancement (PMSO) directing calculation to build, recoup and select disjoint ways that endure the disappointment while fulfilling Quality of Service (QoS) parameters. Multi-swarm technique empowers deciding the optimal directions in choosing the multipath directing while at the same time trading messages from all positions in the system. Results demonstrated that the procedure utilizing the qualities of

all individual best data is a substantial technique for the motivations behind enhancing the PMSO execution.

In 2017 Dina Hussein *et al*. [45] have examined a novel administration structure in view of a reasoning approach for dynamic SIoT administrations revelation in smart spaces. That is, reasoning about clients' situational needs, inclinations, and other social angles along with clients' encompassing condition were proposed for creating a file of situation-aware services which coordinates clients' needs. This reasoning approach is then actualized as a proof-of-idea model, specifically Airport Dynamic Social, inside a brilliant airport.

In 2017 Xuan Thang Nguyen *et al.* [46] investigated the issue in detail and propose a model-driven way to optimize an IoT application in regards to its non-practical prerequisites. A source code transformation, that updates the source code with the produced movable parameter values and then executes the compiler to make another binary image of the application. The investigational results showed that non-useful prerequisites, for example, power consumption and reliability can be enhanced generously at the time of optimization.

III. CONTRIBUTION OF WORK

The SIoT structure can be framed as required to ensure the system safety, so as that the disclosure of objects as well as services is performed effectively and the scalability is ensured like in the human social systems. In paper focus to SIOT big data with data classification model, initially the proposed model considers the SIoT data to remove the noise and redundant data filtering approach will help then Hadoop MapReduce (HMR) programming paradigm used to divides the input database into small independent groups which processed in a parallel manner. From the optimal rules to a classification of the types of social relationships that can be established between objects using Supervised Learning Classifier that is Linear Kernel Support Vector Machine [46-50]. For the purpose of proposed classifier predicting the class of objects whose class label is unknown. This work will be useful for the SIoT systems and will provide the guidance for the proposed research in SIoT and big data.

4.1 SIoT Data collection layer

The developed things in the SIoT can detect the physical condition, collect data, exchange or scatter information, process information for fitting applications, and communicate with different things. Henceforth, SIoT came up with power innovation that aides in understanding the physical world and to reaction to external jolts [51-53].

4.2 Filtering Layer

This is the key to this approach since, for a particular dataset, utilizing some filter which varies the chosen subset of highlights and, subsequently, the execution to remove noise and undesirable data Gabor channel were proposed.

Gabor Filter

Gabor is Gaussian function revised by complex sinusoids. Under specific conditions, the period of the reaction of Gabor channels is roughly linear. Gabor wavelet symbolizes a class of this sort of functions. The utilization of the piece of Gabor channel bank can viably diminish the calculation and decrease the measurement, even enhance the remove noisy information. A Gabor base is a Gaussian function adjusted with an exponential that is characterized as far as the result of a Gaussian and an exponential. Two-dimensional Gabor functions $u(a,b)$ it's by

$$u(a,b) = g(a,b) * \exp^{-2\pi j f_y a} \quad \text{..... (1)}$$

Here $$g(a,b) = \frac{1}{2\pi\sigma_a\sigma_b} * \exp^{\frac{-1}{2}\left(\frac{a^2}{\sigma_a^2}+\frac{b^2}{\sigma_b^2}\right)}$$

..... (2)

Gabor capacities are bandpass filters which are Gaussians focused on frequency in the spatial-area. Gabor channels with various frequencies and with introductions in various ways have been utilized to localize and extricate text area from data by removing noise.

4.3 Database Reduction layer

Hadoop MapReduce (HRM) comprises of a Job Tracker and a few Task Trackers. At the point, when a MapReduce task is executed, the Job Tracker divides it into smaller tasks (map as well as reduce) processed by the task trackers. At the point, when the input dataset is given to a MapReduce data it is split into autonomous data chunks which are, at that point prepared by the map tasks in parallel. The map tasks outputs are given to the input of reduce tasks. The map function executes the technique of allocating each sample to the nearest focus while the reduce tasks executes the new centers updating. From this structure whole data mapped in a few group and the reducer reduce SIoT database based on the threshold value [54-56].

Mapping Process: One set of filtered data is converted into another set of data input. Given the data, a mapper can evaluate the nearest center point for every data. Maps are the individual tasks that change input records into the middle of the road information. The transformed intermediate records shouldn't be of similar type as the input data.

Reducer Process: Secondly, diminish mission, which obtains the productivity from a map as an input and unites those data tuples into a slighter group of tuples and results in a set of values that are reduced in amount. Next is the

reduce process which combines the data created in the mapping process.

IV. SIMULATION RESULTS AND ANALYSIS

This proposed SIoT data classification with Optimal feature selection implemented in Java programming language with JDK 1.7.0 in a windows machine containing configurations 4 GB RAM and the operating system platform.

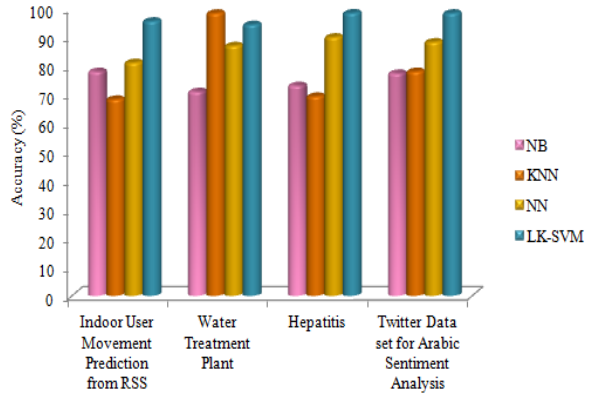

Fig 1: Comparative analysis for Classification Performance

Figure 1 represents the accuracy of the different classifier for the four different databases like Indoor User Movement Prediction from RSS, Water Treatment Plant, Hepatitis and Twitter Dataset for Arabic Sentiment Analysis. For the four databases, the maximum accuracy is accomplished in the proposed LK-SVM compared to other classifiers.

V. CONCLUSION

The study scrutinized the classification of Big data based Social Internet of Things (SIoT) with big data map-reduce system. Here, Gabor filter was utilized to reduce noise and the redundant data from the database. Further to improve the efficiency of the proposed work, Hadoop MapReduce was analyzed to mapping the data. Before the data classification, the optimal features were selected from the database by the inspired algorithms. In this research work, the Elephant Herding Optimization (EHO) was utilized to choose the optimal features from each database. Finally, Linear Kernel Support Vector Machine (LK-SVM) classifier was proposed to classify the data as members or non-members of the class based on the expression data. From the results, the proposed LK-SVM with EHO attains a maximum accuracy of 98.86% compared to other existing approaches. In future, the efficiency i.e. the performance measures of big data classification based SIoT is improved by the utilization of advanced classifier with the hybrid optimization algorithms.

Applications

Some application of the Internet of things in a social network that includes a smart device that can push its status messages on twitter. Also in few sectors like telecommunication, retail and finance have been early adopters of big data analytics.

REFERENCE

[1] Chen, Y. T., Chen, C. H., Wu, S., & Lo, C. C. (2019). A two-step approach for classifying music genre on the strength of AHP weighted musical features. Mathematics, 7(1), 19.

[2] Elhoseny, M., Shankar, K., &Uthayakumar, J. (2019). Intelligent diagnostic prediction and classification system for chronic kidney disease. Scientific reports, 9(1), 1-14.

[3] Sivakumar, P., Velmurugan, S. P., & Sampson, J. (2020). Implementation of differential evolution algorithm to perform image fusion for identifying brain tumor.

[4] Khamparia, A., Gupta, D., Nguyen, N. G., Khanna, A., Pandey, B., & Tiwari, P. (2019). Sound classification using convolutional neural network and tensor deep stacking network. IEEE Access, 7, 7717-7727.

[5] Jansirani, A., Rajesh, R., Balasubramanian, R., & Eswaran, P. (2011). Hi-tech authentication for pslette images using digital signature and data hiding. Int. Arab J. Inf. Technol., 8(2), 117-123.

[6] Jain, R., Gupta, D., & Khanna, A. (2019). Usability feature optimization using MWOA. In International conference on innovative computing and communications (pp. 453-462). Springer, Singapore.

[7] Shankar, K., & Lakshmanaprabu, S. K. (2018). Optimal key based homomorphic encryption for color image security aid of ant lion optimization algorithm. International Journal of Engineering & Technology, 7(9), 22-27.

[8] Lyu, L., & Chen, C. H. (2020, July). Differentially Private Knowledge Distillation for Mobile Analytics. In Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 1809-1812).

[9] Poonkuntran, S., Rajesh, R. S., & Eswaran, P. (2011). Analysis of difference expanding method for medical image watermarking. In International Symposium on Computing, Communication, and Control (ISCCC 2009) (Vol. 1, pp. 31-34).

[10] Sampson, J., & Velmurugan, S. P. (2020, March). Analysis of GAA SNTFT with Different Dielectric Materials. In 2020 5th International Conference on

Devices, Circuits and Systems (ICDCS) (pp. 283-285). IEEE.

[11] Elhoseny, M., Bian, G. B., Lakshmanaprabu, S. K., Shankar, K., Singh, A. K., & Wu, W. (2019). Effective features to classify ovarian cancer data in internet of medical things. Computer Networks, 159, 147-156.

[12] Gochhayat, S. P., Kaliyar, P., Conti, M., Tiwari, P., Prasath, V. B. S., Gupta, D., & Khanna, A. (2019). LISA: Lightweight context-aware IoT service architecture. Journal of cleaner production, 212, 1345-1356.

[13] Dutta, A. K., Elhoseny, M., Dahiya, V., & Shankar, K. (2020). An efficient hierarchical clustering protocol for multihop Internet of vehicles communication. Transactions on Emerging Telecommunications Technologies, 31(5), e3690.

[14] Anand Nayyar, Vikram Puri, Nhu Gia Nguyen, BioSenHealth 1.0: A Novel Internet of Medical Things (IoMT) Based Patient Health Monitoring System, Lecture Notes in Networks and Systems. Springer, 2019

[15] Shankar, K., Lakshmanaprabu, S. K., Khanna, A., Tanwar, S., Rodrigues, J. J., & Roy, N. R. (2019). Alzheimer detection using Group Grey Wolf Optimization based features with convolutional classifier. Computers & Electrical Engineering, 77, 230-243.

[16] Paramathma, M. K., Pravin, A. C., Rajarajan, R., & Velmurugan, S. P. (2019, April). Development and Implementation of Efficient Water and Energy Management System for Indian Villages. In 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS) (pp. 1-4). IEEE.

[17] Chen, C. H., Song, F., Hwang, F. J., & Wu, L. (2020). A probability density function generator based on neural networks. Physica A: Statistical Mechanics and its Applications, 541, 123344.

[18] Kathiresan, S., Sait, A. R. W., Gupta, D., Lakshmanaprabu, S. K., Khanna, A., & Pandey, H. M. (2020). Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model. Pattern Recognition Letters.

[19] Gupta, D., & Ahlawat, A. K. (2016). Usability determination using multistage fuzzy system. Procedia Comput Sci, 78, 263-270.

[20] Amira S. Ashour, Samsad Beagum, Nilanjan Dey, Ahmed S. Ashour, Dimitra Sifaki Pistolla, Gia Nhu Nguyen, Dac-Nhuong Le, Fuqian Shi (2018), Light Microscopy Image De-noising using Optimized LPA-ICI Filter, Neural Computing and Applications, Vol.29(12), pp 1517–1533, Springer, ISSN: 0941-0643. (SCIE IF 4.664, Q1)

[21] Pan, M., Liu, Y., Cao, J., Li, Y., Li, C., & Chen, C. H. (2020). Visual Recognition Based on Deep Learning for Navigation Mark Classification. IEEE Access, 8, 32767-32775.

[22] Chen, C. H., Hwang, F. J., & Kung, H. Y. (2019). Travel time prediction system based on data clustering for waste collection vehicles. IEICE TRANSACTIONS on Information and Systems, 102(7), 1374-1383.

[23] Shankar, K., & Eswaran, P. (2015). ECC based image encryption scheme with aid of optimization technique using differential evolution algorithm. Int J Appl Eng Res, 10(55), 1841-5.

[24] Anand Nayyar, Vikram Puri, Nhu Gia Nguyen, Dac Nhuong Le, Smart Surveillance Robot for the Real Time Monitoring and Control System in Environment and Industrial Applications, Advances in Intelligent System and Computing, pp 229-243, Springer

[25] Le Nguyen Bao, Dac-Nhuong Le, Gia Nhu Nguyen, Vikrant Bhateja, Suresh Chandra Satapathy (2017), Optimizing Feature Selection in Video-based Recognition using Max-Min Ant System for the Online Video Contextual Advertisement User-Oriented System, Journal of Computational Science, Elsevier ISSN: 1877-7503. Vol.21, pp.361-370. (SCIE IF 2.502, Q1)

[26] Chakchai So-In, Tri Gia Nguyen, Gia Nhu Nguyen: Barrier Coverage Deployment Algorithms for Mobile Sensor Networks. Journal of Internet Technology 12/2017; 18(7):1689-1699.

[27] Le, D.-N.a, Kumar, R.b, Nguyen, G.N., Chatterjee, J.M.d, Cloud Computing and Virtualization, DOI: 10.1002/9781119488149, Wiley.

[28] Bhateja, V., Gautam, A., Tiwari, A., Nhu, N.G., Le, D.-N, Haralick features-based classification of mammograms using SVM, Advances in Intelligent Systems and Computing, Volume 672, 2018, Pages 787-795.

[29] Khamparia, A., Saini, G., Gupta, D., Khanna, A., Tiwari, S., & de Albuquerque, V. H. C. (2020). Seasonal crops disease prediction and classification using deep convolutional encoder network. Circuits, Systems, and Signal Processing, 39(2), 818-836.

[30] Uthayakumar, J., Elhoseny, M., & Shankar, K. (2020). Highly Reliable and Low-Complexity Image Compression Scheme Using Neighborhood Correlation Sequence Algorithm in WSN. IEEE Transactions on Reliability.

[31] Huyen, D.T.T., Binh, N.T., Tuan, T.M., Nguyen, G.N, Dey, N., Son, L.H, Analyzing trends in

hospital-cost payments of patients using ARIMA and GIS: Case study at the Hanoi Medical University Hospital, Vietnam, Journal of Medical Imaging and Health Informatics, 7(2), pp. 421-429.

[32] Van, V.N., Chi, L.M., Long, N.Q., Nguyen, G.N., Le, D.-N, A performance analysis of openstack open-source solution for IaaS cloud computing, Advances in Intelligent Systems and Computing, 380, pp. 141-150.

[33] Shankar, K., & Eswaran, P. (2016, January). A new k out of n secret image sharing scheme in visual cryptography. In 2016 10th International Conference on Intelligent Systems and Control (ISCO) (pp. 1-6). IEEE.

[34] Dey, N., Ashour, A.S., Chakraborty, S., Le, D.-N., Nguyen, G.N, Healthy and unhealthy rat hippocampus cells classification: A neural based automated system for Alzheimer disease classification, Journal of Advanced Microscopy Research, 11(1), pp. 1-10

[35] Velmurugan, S. P., & Rajasekaran, P. S. M. P. (2017). CLASSIFICATION OF BRAIN TUMOR USING MULTIMODAL FUSED IMAGES AND PNN. International Journal of Pure and Applied Mathematics, 115(6), 447-457.

[36] Shankar, K., Elhoseny, M., Perumal, E., Ilayaraja, M., & Kumar, K. S. (2019). An Efficient Image Encryption Scheme Based on Signcryption Technique with Adaptive Elephant Herding Optimization. In Cybersecurity and Secure Information Systems (pp. 31-42). Springer, Cham.

[37] Wu, L., Chen, C. H., & Zhang, Q. (2019). A mobile positioning method based on deep learning techniques. Electronics, 8(1), 59.

[38] Lydia, E. L., Kumar, P. K., Shankar, K., Lakshmanaprabu, S. K., Vidhyavathi, R. M., & Maseleno, A. (2020). Charismatic document clustering through novel K-Means non-negative matrix factorization (KNMF) algorithm using key phrase extraction. International Journal of Parallel Programming, 48(3), 496-514.

[39] Sujitha, B., Parvathy, V. S., Lydia, E. L., Rani, P., Polkowski, Z., & Shankar, K. (2020). Optimal deep learning based image compression technique for data transmission on industrial Internet of things applications. Transactions on Emerging Telecommunications Technologies, e3976.

[40] Lo, C. L., Chen, C. H., Hu, J. L., Lo, K. R., & Cho, H. J. (2019). A fuel-efficient route plan method based on game theory. Journal of Internet Technology, 20(3), 925-932.

[41] Kung, H. Y., Chen, C. H., Lin, M. H., & Wu, T. Y. (2019). Design of Seamless Handoff Control Based on Vehicular Streaming Communications. Journal of Internet Technology, 20(7), 2083-2097.

[42] Elhoseny, M., & Shankar, K. (2019). Reliable data transmission model for mobile ad hoc network using signcryption technique. IEEE Transactions on Reliability.

[43] Shanmugam, P., Rajesh, R. S., & Perumal, E. (2008, May). A reversible watermarking with low warping: an application to digital fundus image. In 2008 International Conference on Computer and Communication Engineering (pp. 472-477). IEEE.

[44] Shankar, K., & Elhoseny, M. (2019). Trust Based Cluster Head Election of Secure Message Transmission in MANET Using Multi Secure Protocol with TDES. J. UCS, 25(10), 1221-1239.

[45] Parvathy, V. S., Pothiraj, S., & Sampson, J. (2020). Optimal Deep Neural Network model based multimodality fused medical image classification. Physical Communication, 101119.

[46] Subbiah Parvathy, V., Pothiraj, S., & Sampson, J. (2020). A novel approach in multimodality medical image fusion using optimal shearlet and deep learning. International Journal of Imaging Systems and Technology.

[47] Parvathy, V. S., & Pothiraj, S. (2019). Multi-modality medical image fusion using hybridization of binary crow search optimization. Health Care Management Science, 1-9.

[48] Velmurugan, S. P., Sivakumar, P., & Rajasekaran, M. P. (2018). Multimodality image fusion using centre-based genetic algorithm and fuzzy logic. International Journal of Biomedical Engineering and Technology, 28(4), 322-348.

[49] Chen, C. H. (2018). An arrival time prediction method for bus system. IEEE Internet of Things Journal, 5(5), 4231-4232.

[50] Shankar, K., Perumal, E., & Vidhyavathi, R. M. (2020). Deep neural network with moth search optimization algorithm based detection and classification of diabetic retinopathy images. SN Applied Sciences, 2(4), 1-10.

[51] Mohanty, S. N., Ramya, K. C., Rani, S. S., Gupta, D., Shankar, K., Lakshmanaprabu, S. K., & Khanna, A. (2020). An efficient Lightweight integrated Blockchain (ELIB) model for IoT security and privacy. Future Generation Computer Systems, 102, 1027-1037.

[52] Elhoseny, M., & Shankar, K. (2020). Energy efficient optimal routing for communication in VANETs via clustering model. In Emerging Technologies for Connected Internet of Vehicles and Intelligent Transportation System Networks (pp. 1-14). Springer, Cham.

[53] Chen, C. H. (2020). A cell probe-based method for vehicle speed estimation. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 103(1), 265-267.

[54] Furqan Alama, Rashid Mehmood, Iyad Katiba and Aiiad Albeshri,"Analysis of Eight Data Mining Algorithms for Smarter Internet of Things (IoT)", Procedia Computer Science,V ol.98,pp. 437-442,2016.

[55] Khamparia, A., Singh, A., Anand, D., Gupta, D., Khanna, A., Kumar, N. A., & Tan, J. (2018). A novel deep learning-based multi-model ensemble method for the prediction of neuromuscular disorders. Neural computing and applications, 1-13.

[56] Shankar, K., Zhang, Y., Liu, Y., Wu, L., & Chen, C. H. (2020). Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification. IEEE Access, 8, 118164-118173.

AUTHOR DETAILS:

Dontala kiran kumar
M.tech (cse) baba institute of technology and sciences (bits vizag) bakkannapalem, pm palem, madhurawada, visakhapatnam.

Suragani.chanti
M.tech (cse)
Asst professor baba institute of technology and sciences (bits vizag) bakkannapalem, pm palem, madhurawada, visakhapatnam.

Kaligiti satyanarayana murthy
Btech, mtech, (phd)
Asst professor and project co-ordinator baba institute of technology and sciences (bits vizag) bakkannapalem, pm palem, madhurawada, visakhapatnam.