Different analytical frameworks and bigdata model for internet of things

Ayushi Chahal, Preeti Gulia, Nasib Singh Gill
Department of Computer Science and Applications, Maharshi Dayanand University, Rohtak, Haryana, India

Article Info

Article history:

Received Mar 26, 2021
Revised Dec 15, 2021
Accepted Dec 22, 2021

Keywords:

Bigdata
Big IoT data analytics
Frameworks
IoT
Machine learning

ABSTRACT

Sensor devices used in internet of things (IoT) enabled environment produce large amount of data. This data plays a major role in bigdata landscape. In recent years, correlation, and implementation of bigdata and IoT is being extrapolated. Nowadays, predictive analytics is gaining attention of many researchers for big IoT data analytics. This paper summarizes different sort of IoT analytical platforms which consist in-built features for further use in machine learning, MATLAB, and data security. It emphasizes on different machine learning algorithms that plays important role in big IoT data analytics. Besides different analytical frameworks, this paper highlights the proposed model for bigdata in IoT domain and elaborates different forms of data analytical methods. Proposed model comprises different phases i.e., data storing, data cleaning, data analytics, and data visualization. These phases cover the basic characteristics of bigdata V's model and most important phase is data analytics or big IoT analytics. This model is implemented using an IoT dataset and results are presented in graphical and tabular form using different machine learning techniques. This study enhances researchers' knowledge about various IoT analytical platforms and usability of these platforms in their respective problem domains.

Corresponding Author:

Ayushi Chahal
Department of Computer Science and Applications, Maharshi Dayanand University
Rohtak, Haryana, India
Email: ayushichahal@gmail.com

1. INTRODUCTION

In today's era, internet of things (IoT) and bigdata are potential areas of research. Heterogeneous data is generated by the different sensor's devices in the IoT environment. Bigdata analytics is used to search, mine, and analyze IoT data. It can also be used to handle structured, semi-structured and unstructured data [1] and helps to convert this data into some understandable form for the analysis. There are multiple techniques that can be used for bigdata analytics such as classification, clustering, association rules, and prediction. Bigdata is characterized by 10 V's: Volume, Velocity, Variety, Value, Veracity, Validity, Variability, Viscosity, Virality and Visualization [2] shown in Figure 1.

Gartner defines bigdata concept that helps in decision making, optimizing the processes, discover patterns insightfully. He gave a characteristics model for bigdata which defines three V's that is volume, velocity and variety of data. Gartner research has made estimation that by 2022 most of the data generation and its analysis will be done by machines rather than humans. So, it is need of an hour to have a model which can handle Big IoT data efficiently for prediction using Machine learning techniques.

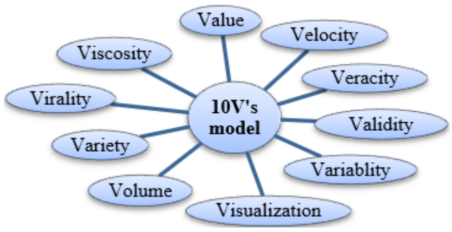

Figure 1. The 10 V's of bigdata

L'Heureux *et al.* [3] identified research opportunities by combining these two new technologies. Some researchers like Al-Jarrah *et al.* [4], Najafabadi *et al.* [5], Sukumar [6], and Qiuet *et al.* [7] also explored the challenges of machine learning that one has to face while dealing with bigdata. Marjani *et al*. [8] enable people to have a perception about large data generated by different sensor devices, in different commercial sectors and also explained how to deal and analyse this large data using different machine learning techniques like classification, clustering, Prediction and association rule mining. It explained how IoT is related to bigdata analytics.

Habibzadeh *et al*. [9] helps us understanding the smart cities sensing networks. They have also described how machine intelligence and data analytics algorithms are used in such a scenario. They described various smart city applications in their paper. Intelligent transportation system, health monitoring, smart parking, intelligent lighting, smart grid, smart utilities etc are explained in the study. Transportation is said to be smart, when smart roads automatically notify the driver about the bad traffic conditions. Parking is said to be smart, if parking space communicates with the drivers about the location of unoccupied parking spaces. An environment is assumed to be smart when it empowers smart homes and smart workplaces to balance their temperature to conserve energy. The paper tries to explain corresponding machine learning as well as data analytics algorithms, for every application.

Kaur and Kushwaha [10] have also explained nicely the role of bigdata in IoT analytics and also have focused on integration of these two vast technologies. He explained different platforms for bigdata analytics like Apache Hadoop, spark, map-reduce, 1010 data, hp-hive etc. These platforms that can be used for IoT data sets too. They also discussed bigdata taxonomy with IoT analytical solutions.

Sagu *et al*. [11] explained IoT is being used worldwide and hence needed to be secured. For this security purpose, authors have elaborated different ways of artificial neural networks to be used. Ratra and Gulia [12] have described different data mining techniques that can be used for security purpose. Al-Shorman *et al*. [13] described healthcare domain and use of IoT data analytics in it. For this description authors have used real time data processing for diabetic patient's case study. Rahman *et al*. [14] discussed different security issues of using IoT with mobiles with the help of edge computing.

In the era of highly connected network, there is an explosion of information flow with the help of IoT. Since past few years, this flow of information has gained such a momentum which ensures that, in future, connectivity between different gadgets and devices will be handled by internet of everything (IoE). Such abundant heterogeneous data give rise to a new concept called "big IoT data". This big IoT [15] data needs to be analysed which helps in improving understandability of raw data, so that efficient and well-informed decisions can be made [16], [17].

Big IoT data can be available in different forms from different smart devices, such as: unstructured data format, fast moving data (in streaming form), noisy and poor-quality data, highly dimensional data, imbalanced distributed data, unlabelled data, limited labelled data. These are the obstacles which are needed to overcome before applying any kind of data analytics algorithm [18]. Nicolalde *et al.* [19] discuss integration of bigdata analytics and IoT with their challenges like knowledge discovery and computational complexities, data storage, and information security. They have also referred different tools to overcome these challenges. Different big IoT data analytical techniques are shown in Figure 2.

- Descriptive analytics focuses on "what". This analytics is done when data is accumulated i.e., at first step. Basic nature and certain patterns can be easily found out with the help of descriptive analytics. It uses most common data mining techniques like classification, clustering and segmentation of data.
- Diagnostic analytics focuses on "why". It mainly depends on the incidents that have appeared in past. Both descriptive and diagnostic analysis can't predict the futuristic behavior. It relays on machine learning algorithms to answer "why" questions to the data.
- Predictive analytics is used for prediction based forecast. Predictive analytics helps to find future patterns that will occur with the help of present data. It generally uses machine learning and statistical algorithms for future prediction.

- Perspective analytics is relative to both descriptive as well as diagnostic analytics. It can answer all the questions such as "what", "when", "why", "how" should be done. Perspective analytics uses Artificial Intelligence algorithms to process.

Bigdata analytics uses machine learning for implementing various data mining techniques [19]. These techniques help to take out valuable information from raw big IoT data. This information helps in predicting results, in decision making, identifying patterns and trends, and discovering hidden information [20]. Large amount of IoT data is processed, transformed, and analyzed at high frequency [21]. Various solutions are present in the market for bigdata analytics.

- Machine learning is assumed to be the most fundamental component for big IoT data analytics [22]. On basis of their learning tasks, the data learning process is categorized as: supervised and unsupervised learning.
- Supervised learning is learning, in which both inputs and corresponding outputs are already known to the system. With this knowledge the system learns to map inputs data to output data for a particular system. Classification and regression are supervised learning techniques. In classification supervised learning method, discrete values are taken by outputs. Examples of classification algorithm are: k-nearest neighbor, Naïve, logistic regression and support vector machine (SVM).
- Unsupervised learning is machine learning technique, in which desired outputs for the data are unknown to the system. In unsupervised learning, system itself tries to find out the patterns within the data . It includes clustering method. In clustering method, grouping of data points or objects is done on the basis of sone sort of similarity criteria. One of the examples of clustering method is k-means algorithm.
- Hadoop and Spark are tools for bigdata analytics that can be used in the healthcare and transportation domains [23]-[25] very efficiently. MapReduce is used for parallel computing and distributed storage IoT environment [26].
- Deep learning [27], fuzzy logic [28], data envelopment analysis [29] are some of the variants of machine learning which came out be an effective variant of machine learning.

Researchers need to focus on synchronization between bigdata and different analytical techniques so that they can make easier and better-informed choices in taking decisions. This study focuses on the relationship of bigdata with IoT and this relationship is used to propose a model for big IoT data analysis. A full article usually follows a standard structure: i) Introduction, ii) IoT analytics Frameworks, iii) Proposed model of IoT bigdata, iv) Results and Discussion, and v) Conclusion. Different IoT analytics frameworks are discussed in section 2, which will help researchers to finalize what kind of framework they should opt with respect to their dataset. Section 3 elaborates details of the proposed model with different phases. This proposed model is applied on a dataset named "Biochemical features of orthopedic patient" as discussed in section 4 and found promising results.

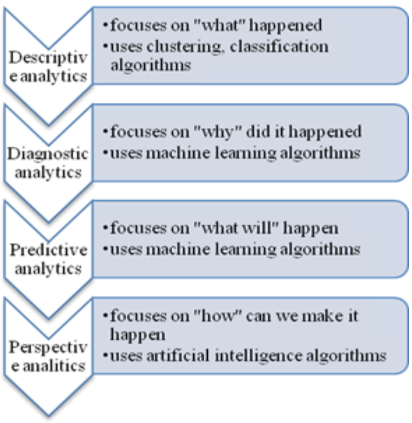

Figure 2. Different big IoT data analytical techniques

2. IOT ANALYTICS FRAMEWORKS

An IoT data analytics platform must have ability to manage gigabytes of data generated by different IoT devices connected over the network. IoT data analytics platform must be able to ingest structured, semi-structured, unstructured, real time or time series, sequential data so that intelligent decisions can be taken. Some of these frameworks are explained below and remaining is described briefly in Table 1.

Table 1. Different IoT analytical framework and their description

No.	IoT Frameworks	Framework Symbol	Framework Description
1.	Google Cloud IoT Core		Provides secure connections between IoT devices Deploy machine learning models Integrates Google Bigdata analytics and ML services
2.	Thing Speak		Helps in real time sensor data collection Provides data analytics and visualization with the help of built-in MATLAB data analytics tools Works with Arduino and Raspberry Pi.
3.	AWS IoT Analytics	aws	Can perform analytics on massive volumes of IoT data One can build their own IoT analytics platform without worrying about cost built in tools such as MATLAB and Octave.
4.	AT&T IoT platform		It provides cloud services for storing IoT data and connectivity between IoT devices Provide data Visualization, Global Connectivity & Management, Data Orchestration
5.	Oracle Internet of Things Cloud	ORACLE CLOUD	Provides end to end security Uses IoT SaaS applications Have in–build ML capabilities
6.	Aptche incubator		Apache iota is sponsored by the Incubator. It is an effort undergoing incubation at Apache Software Foundation (ASF). But now Iota project is retired
7.	AWS IoTSiteWise	aws	Helps in collecting, storing, organizing, and monitoring the data in a very easy manner With remote monitoring it identifies issues very quickly Uses central data source to improve cross-facility process
8.	Bosch IoT Suite	BOSCH	Enables users to connect directly or indirectly to the devices via clouds or gateways respectively It helps in storing and updating the data, properties, and relations of the user, which in turn helps in keeping the real and digital worlds synchronized.
9.	SAS Analytics for IoT	sas	Uses Sensor based data model Have advanced analytics and embedded AI Analyze data without writing code It applies various machine learning algorithms
10.	SAP Leonardo Internet of Things	SAP	Cloud deployment It upgrades IoT data to rich business context It provides various Analytical services and query model for its application
11.	IBM Watson IoT	IBM	Securely connect, manage, and analyze IoT data. It provides data storage for IoT and its rapid visualization.
12.	Knowi	knowi	One can Prepare their own Training Data Evaluate and Train Models Integrate Model into Any Analytics Workflow Trigger Data-Driven Actions
13.	Ubidots		It easily captures sensor data and turn it into useful information. It transforms raw data into understandable format with Synthetic Variables It can compute complex mathematical formulas and statistical expressions.
14.	Thing+	Thing+	It offers the tools for reporting and investigating data coming out of IOT devices. Provides data comparisons. Offers historical analysis. Helps in finding trends.
15.	Microsoft Azure	Microsoft Azure	Azure IoT Hub provides a cloud-hosted solution manage IoT devices at scale

a. Amazon Web Service IoT (AWS IoT): It allows a secure communication among different IoT devices which are connected with each other. It helps to store data over the cloud and helps to analyse Peta Bytes of data which helps in IoT data analytics. AWS IoT works over four layers namely: device gateway, rules engine, device shadows and registry [30].
b. Iotivity: It is a freely available IoT data analytics framework. It is basically used in Smart Home Fields. Iotivity has a layered architecture which consists of three layers: the base layer, the service layer, and the cloud interface. Base layer is a very first layer, which helps in different device connectivity with maximum security. Service layer provides the simulator which helps in testing devices before purchasing them.
c. Azure IoT Suite: It is Microsoft released IoT framework and not an open source. It helps different IoT devices to remain connected with the help of cloud. It provides virtual environment for testing devices. It helps in machine learning and real time data analytics. Its main part is Azure IoT Hub which helps in

providing security, authenticity, and connection between personal area network (PAN) via Gateways. It provides data visualization.

d. Eclipse Kura: It is a Java based freely available framework. It supports Linux based devices only. It does not provide data visualization. It does not support machine learning algorithms. It can be used for data cleaning but not for data analytics. It provides good storage and security to the data for machine-to-machine applications with the help of gateways [31].

e. Smart Things: It is a paid platform given by Samsung for IoT. It can be used for Smart home applications to control, monitor them via smart phones. It provides storage and security with authentication control. It uses HTTP as a messaging protocol for machine to machine data passing. Groovy is used as a programming environment for smart thing hub [32].

3. PROPOSED MODEL FOR IOT BIGDATA

Bigdata is very important for observing hidden patterns and behavior of the system bigdata can do so, with the help of data collected by the system over long period of time. One can reach to different conclusions based on different functions that can be applied on bigdata. These functions are: aggregation. cleaning, combining, gathering, modelling, selection, annotation, compression, extraction, indexing, mining, storage, analytics, clustering, evaluation, integration, recording, transformation, prediction, transportation, representation, replication, retrieval, searching, stream processing, selection, storage, and visualization [33]. But for big IoT data, there is no need to perform all these functions to reach any conclusion. The proposed model covers characteristics of bigdata as well as helps in analysis of IoT data too. These four phases are described below in Figure 3.

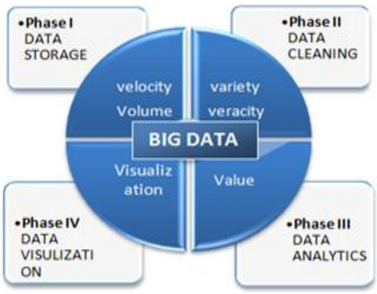

Figure 3. Phases of bigdata in IoT domain

- Phase 1: Data storage
  It deals with two characteristics of bigdata i.e. velocity and volume. As bigdata is a bulk of data moving really fast. For storing IoT data cloud is most favorable technology. It can store data for NoSQL as well as relational databases. NoSQL is used more commonly because it can hold unstructured or semi-structured data very easily.
- Phase 2: Data cleaning or data cleansing
  It deals with another bigdata categories i.e. variety and veracity. bigdata is produced by heterogeneous sources with multiple trait levels. Data cleaning basically works on two aspects i.e. data integration and data quality management. Data integration is another form of extract, transform, load (ETL). Data quality management deals with corrupted data detection, data redundancy reduction, data integration check. In next phase i.e data analytics, quality depends on the type of data send to it after cleaning.
- Phase 3: Data analytics or data analysis
  It helps to came out with value driven outputs with insightful and valuable interpretations. bigdata analytics or analysis techniques can widely be used in IoT domain. Clustering is one of the most common algorithm that is used in every sector for IoT data analytics either it is healthcare, transport, and agriculture or energy.
- Phase 4: Data visualization
  It creates the view for the outcomes of data analytics/analysis phase. In it data is represented and interpreted in different ways by different methods. Some-times it is also called as data interpretation or presentation. Machine learning can help in improving data visualization i.e. functioning, reliability and scalability. Data visualization helps in improving analytics by visuals.

4. RESULT AND DISCUSSION

Practical implementation of the proposed model is explained on a dataset named “Biochemical features of orthopedic patient” [34] from kaggle. This dataset contains 310 records with six biomechanical attributes. These attributes are taken according to the orientation of pelvis and lumbar spine. For implementation of this model the chosen dataset is already in the cleaned form. For the analysis of “Biochemical features of orthopedic patient” dataset, different supervised machine learning algorithms [35] like k-nearest neighbors (KNN), logical regression (LR), decision tree (DT), support vector machine (SVM), Naïve Bayes (NB), and Random forest (RF) [36] are used. The dataset is used to analyze and predict, if the patient belongs to normal or abnormal category.

Accuracy and confusion matrix are used as a measure to evaluate above said machine learning algorithms. This analysis is done using anaconda Jupyter notebook. Different libraries of python like numpy, pandas, and sklearn are used for the implementation of these machine learning algorithms. After analysis of different machine learning algorithms results are displayed in Table 2. After analysis of dataset and its visualization is done by importing libraries of Python like matplotlib and seaborn are shown in Figure 4 and Figure 5. Graphic view of the accuracy ploted using KNN and SVM machine learning algorithms are shown in Figure 6 and Figure 7 respectively. A comparative analysis of variable k in KNN algorithm and acuraccy is presented in Figure 8. Overall performance analysis of machine learning algorithm is depicted through confusion matrix in Figure 9.

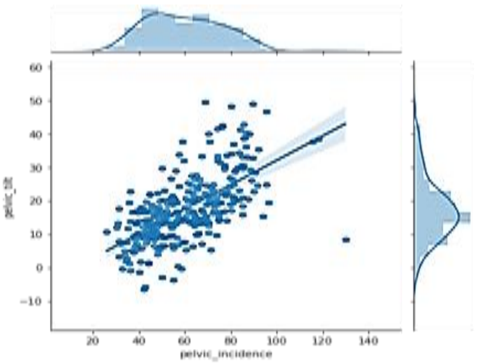

Figure 4. Joint plot of dataset

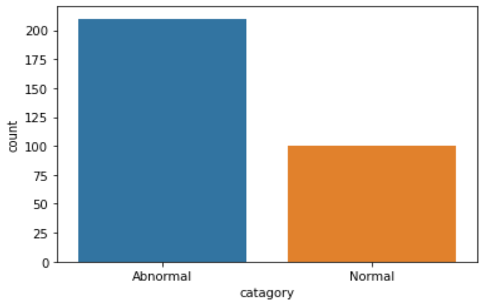

Figure 5. Count-plot of category

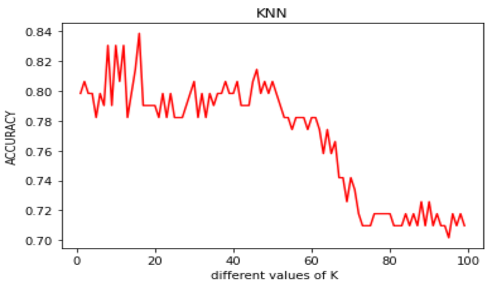

Figure 6. Accuracy plot for KNN

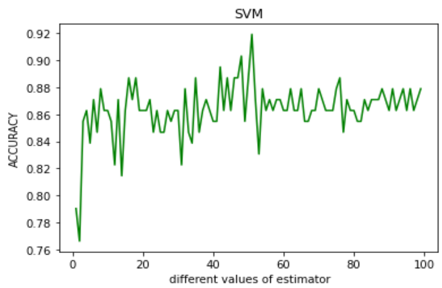

Figure 7. Accuracy plot for SVM

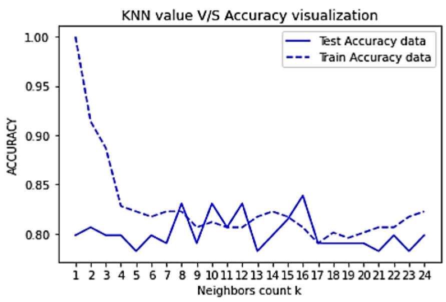

Figure 8. Value of k V/S accuracy

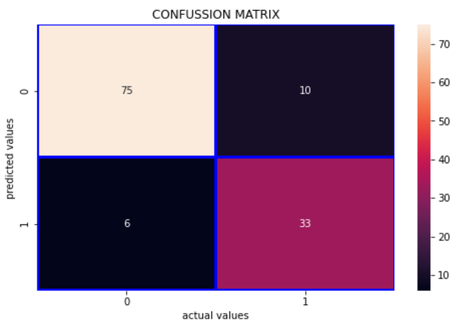

Figure 9. Confusion matrix

Table 2. Score values of different ML algorithms

Diffrerent algorithms	KNN	Logical regression	Decision Tree	SVM	Naïve Bayes	Random forest
Score value	0.8306	0.7258	0.8709	0.8145	0.7580	0.9032

5. CONCLUSION

It is believed that in today's world, each, and every individual is carrying a sensor containing device which generates and receives abundant amount of data. This data is useless if, it is not used for analysis. This study has presented the relationship between IoT and the data analytics. Some of these real-world examples are highlighted in this paper. Different forms of data analytics (i.e., descriptive, diagnostic, predictive, and perspective) and different kind of methods that can be used to execute these analytical methods are explained. It may be concluded that creation of strategy for clustering on big IoT data is a significant challenge for the analyst in big IoT data analytics. Different IoT analytical frameworks with their features are presented in this paper. Deep study on these frameworks may help researchers to find suitable platform according to their needs. The proposed model of bigdata for IoT domain is explained in this study which covers the basic features of IoT and bigdata. Implementation of this model is shown with the help of an IoT dataset in which predictive analytics and data visualization is very well elaborated. In future, this proposed model will be implemented for a big IoT data set to get the best results and comparison of this model will be done with other models too.

REFERENCES

[1] L. Kosan, "Top IoT Data Analytics Platforms," (2019). Accessed: Dec. 14, 2021. [Online]. Available: https://www.iotworldtoday.com/2019/07/10/top-iot-data-analytics-platforms/

[2] A. Ilapakurti, J. S. Vuppalapati, S. Kedari, S. Kedari, C. Chauhan and C. Vuppalapati, "iDispenser-Bigdata Enabled Intelligent Dispenser," *2017 IEEE Third International Conference on Bigdata Computing Service and Applications (BigDataService)*, 2017, pp. 124-130, doi: 10.1109/BigDataService.2017.53.

[3] A. L'Heureux, K. Grolinger, H. F. Elyamany and M. A. M. Capretz, "Machine Learning With Bigdata: Challenges and Approaches," in *IEEE Access*, vol. 5, pp. 7776-7797, 2017, doi: 10.1109/ACCESS.2017.2696365.

[4] O. Y. Al-Jarrah, P. D. Yoo, S. Muhaidat, G. K. Karagiannidis, and K. Taha, "Efficient Machine Learning for Bigdata: A Review," *Bigdata Research*, vol. 2, no. 3, pp. 87-93, Sep. 2015, doi: 10.1016/j.bdr.2015.04.001.

[5] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in bigdata analytics," *Journal of Bigdata*, vol. 2, no. 1, p. 1, Feb. 2015, doi: 10.1186/s40537-014-0007-7.

[6] S. R. Sukumar, "Machine Learning in the Bigdata Era: Are We There Yet?," *Conference: ACM Knowledge Discovery and Data Mining : Workshop on Data Science for Social Good*, Aug. 2014.

[7] J. Qiu, Q. Wu, G. Ding, Y. Xu, and S. Feng, "A survey of machine learning for bigdata processing," *EURASIP Journal on Advances in Signal Processing*, vol. 2016, no. 1, p. 67, May 2016, doi: 10.1186/s13634-016-0355-x.

[8] M. Marjani *et al*., "Big IoT Data Analytics: Architecture, Opportunities, and Open Research Challenges," in *IEEE Access*, vol. 5, pp. 5247-5261, 2017, doi: 10.1109/ACCESS.2017.2689040.

[9] H. Habibzadeh, A. Boggio-Dandry, Z. Qin, T. Soyata, B. Kantarci and H. T. Mouftah, "Soft Sensing in Smart Cities: Handling 3Vs Using Recommender Systems, Machine Intelligence, and Data Analytics," in *IEEE Communications Magazine*, vol. 56, no. 2, pp. 78-86, Feb. 2018, doi: 10.1109/MCOM.2018.1700304.

[10] H. Kaur and A. S. Kushwaha, "A Review on Integration of Bigdata and IoT," *2018 4th International Conference on Computing Sciences (ICCS)*, 2018, pp. 200-203, doi: 10.1109/ICCS.2018.00040.

[11] A. Sagu, N. S. Gill, and P. Gulia, "Artificial Neural Network for the Internet of Things Security" in *International Journal of Engineering Trends and Technology*. vol. 68. pp. 129-136, 2020, doi: 10.14445/22315381/IJETT-V68I11P218.

[12] R. Ratra and P. Gulia, "Privacy Preserving Data Mining: Techniques and Algorithms" in *International Journal of Engineering Trends and Technology*, vol. 68, pp. 56-62, 2020, doi: 10.14445/22315381/IJETT-V68I11P207.

[13] O. Al-Shorman, B. Al-Shorman, M. Al-khassaweneh, and F. Al-Kahtani, "A review of internet of medical things (IoMT) - based remote health monitoring through wearable sensors: a case study for diabetic patients," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 20, no. 1, no. 1, Oct. 2020, doi: 10.11591/ijeecs.v20.i1.pp414-422.

[14] A. Rahman, G. Wu, and A. M. Liton, "Mobile edge computing for internet of things (IoT): security and privacy issues," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 18, no. 3, Art. no. 3, Jun. 2020, doi: 10.11591/ijeecs.v18.i3.pp1486-1493.

[15] T. Sutikno, H. S. Purnama, A. Pamungkas, A. Fadlil, I. M. Alsofyani, and M. H. Jopri, "Internet of things-based photovoltaics parameter monitoring system using NodeMCU ESP8266," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 11, no. 6, pp. 5578-5587, Dec. 2021, doi: 10.11591/ijece.v11i6.pp5578-5587.

[16] A. Chahal, P. Gulia, "Machine Learning and Deep Learning" in *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 8, no. 12, pp. 4910-4914, 2018, doi: 10.35940/ijitee.L3550.1081219.

[17] P. Gulia and A. Chahal, "Bigdata analytics for IoT " in *International Journal of Advanced Research in Engineering and Technology*, vol. 11, no. 6, pp. 593-603, 2020, doi: 10.14445/22315381/IJETT-V68I7P205S.

[18] A. Chahal and P. Gulia, "Deep Learning: A Predictive IoT Data Analytics Method" in *SSRG International Journal of Engineering Trends and Technology (IJETT)*, vol. 68, no. 7, pp. 25-33, 2020, doi: 10.14445/22315381/IJETT-V68I7P205S.

[19] F. Constante Nicolalde, F. Silva, B. Herrera, and A. Pereira, "Bigdata Analytics in IOT: Challenges, Open Research Issues and Tools," in *Trends and Advances in Information Systems and Technologies*, pp. 775–788, 2018, doi: 10.1007/978-3-319-77712-2_73.

[20] D. Pal, T. Triyason, and P. Padungweang, "Bigdata in Smart-Cities: Current Research and Challenges," in *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*, vol. 6, no. 4, Dec. 2018, doi: 10.52549/ijeei.v6i4.543.

[21] B. Karakostas, "Event Prediction in an IoT Environment Using Naïve Bayesian Models," in *Procedia Computer Science*, vol. 83, pp. 11–17, 2016, doi: 10.1016/j.procs.2016.04.093.

[22] J. Betty Jane and E. N. Ganesh, "Bigdata and Internet of Things for Smart Data Analytics Using Machine Learning Techniques," in *Proceeding of the International Conference on Computer Networks, Bigdata and IoT (ICCBI - 2019)*, 2020, pp. 213-223.

[23] Y. Wang, L. Kung, and T. A. Byrd, "Bigdata analytics: Understanding its capabilities and potential benefits for healthcare organizations," *Technological Forecasting and Social Change*, vol. 126, pp. 3–13, 2018, doi: 10.1016/j.techfore.2015.12.019.

[24] M. K. Hassan, A. I. ElDesouky, S. M. Elghamrawy, and A. M. Sarhan, "Bigdata Challenges and Opportunities in Healthcare Informatics and Smart Hospitals," *Security in Smart Cities: Models, Applications, and Challenges*, 2019, pp. 3–26. doi: 10.1007/978-3-030-01560-2_1.

[25] Y. Beeharry, T. P. Fowdur, V. Hurbungs, V. Bassoo, and V. Ramnarain-Seetohul, "Analysing Transportation Data with Open Source Bigdata Analytic Tools," *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*, vol. 5, no. 2, no. 2, Jun. 2017, doi: 10.52549/ijeei.v5i2.297.

[26] D. Dziak, B. Jachimczyk, and W. J. Kulesza, "IoT-Based Information System for Healthcare Application: Design Methodology Approach," *Applied Sciences*, vol. 7, no. 6, no. 6, Jun. 2017, doi: 10.3390/app7060596.

[27] Q. Zhang, L. T. Yang, Z. Chen, and P. Li, "A survey on deep learning for bigdata," *Information Fusion*, vol. 42, pp. 146–157, Jul. 2018, doi: 10.1016/j.inffus.2017.10.006.

[28] D. Zhang, "High-speed train control system bigdata analysis based on the fuzzy RDF model and uncertain reasoning," *Int. J. of Comp. Communications & Control*, vol. 12, no. 4, no. 4, Jun. 2017, doi: 10.15837/ijccc.2017.4.2914.

[29] D. Khezrimotlagh, J. Zhu, W. D. Cook, and M. Toloo, "Data envelopment analysis and bigdata," *European Journal of Operational Research*, vol. 274, no. 3, pp. 1047–1054, May 2019, doi: 10.1016/j.ejor.2018.10.044.

[30] A. Atmani, I. Kandrouch, N. Hmina, and H. Chaoui, "Bigdata for Internet of Things: A Survey on IoT Frameworks and Platforms," *Advanced Intelligent Systems for Sustainable Development (AI2SD'2019)*, pp. 59–67, 2020, doi: 10.1007/978-3-030-33103-0_7.

[31] C. Guindon, "The Eclipse Foundation." Accessed: Dec. 14, 2021. [Online]. Available: https://www.eclipse.org/kura/

[32] SmartThings, "Groovy with SmartThings-SmartThings Classic Developer Documentation." (2021). Accessed: Dec. 15, 2021. [Online]. Available: https://www.smartthings.com/

[33] C. Forrest, "Ten examples of IoT and bigdata working well together." Accessed: Dec. 15, 2021. [Online]. Available: https://www.zdnet.com/article/ten-examples-of-iot-and-big-data-working-well-together/

[34] "Biomechanical features of orthopedic patients." Accessed: Dec. 15, 2021. [Online]. Available: https://kaggle.com/uciml/biomechanical-features-of-orthopedic-patients

[35] M. A. I. Arif, S. I. Sany, F. Sharmin, M. S. Rahman, and M. T. Habib, "Prediction of addiction to drugs and alcohol using machine learning: A case study on Bangladeshi population," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 5, pp. 4471–4480, 2021, doi: 10.11591/ijece.v11i5.pp4471-4480.

[36] R. R. K. AL-Taie, B. J. Saleh, A. Y. F. Saedi, and L. A. Salman, "Analysis of WEKA data mining algorithms Bayes net, random forest, MLP and SMO for heart disease prediction system: A case study in Iraq," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 6, no. 6, pp. 5229-5239, Dec. 2021, doi: 10.11591/ijece.v11i6.

BIOGRAPHIES OF AUTHORS

Ayushi Chahal has passed M.Tech in 2016 in Computer Science and Engineering from Department of Computer Science and Engineering, GJUS&T, Hisar, India. She is currently pursuing Ph.D. in Computer Science at Department of Computer Science and Application, M. D. University, Rohtak, India. Her research interests include IoT, Machine Learning, Bigdata Analytics and Network Security. She can be contacted at email: ayushichahal@gmail.com.

Dr. Preeti Gulia is currently working as Assistant Professor at Department of Computer Science & Applications, M.D. University, Rohtak, India. She is serving the Department since 2009. She earned her doctoral degree in 2013. She has published more than 65 research papers and articles in journal and conferences of National/ International repute including ACM, Scopus. Her area of research includes Data Mining, Bigdata, Machine Learning, Deep Learning, IoT, and Software Engineering. She is an active professional member of IAENG, CSI and ACM. She is also serving as Editorial Board Member Active Reviewer of International/ National Journals. She has guided one research scholar as well as guiding four Ph.D. research scholars from various research areas. She can be contacted at email: preeti.gulia81@gmail.com.

Dr. Nasib Singh Gill is Head of the Department of Computer Science & Applications, M. D. University, Rohtak, India. He is also working as Director, Directorate of Distance Education as well as Director of Digital Learning Centre, M. D. University, Rohtak, Haryana. He earned his Doctorate in Computer Science in the year 1996 and carried out his post-Doctoral research at Brunel University, West London during 2001-2002. He is a recipient of Commonwealth Fellowship Award of British Government for the Year 2001. Besides, he also has earned his MBA degree. He has published more than 304 research papers and authored 5 popular books He has guided so far 12 Ph.D. scholars as well as guiding about 5 more scholars. His research interests primarily includes-IoT, Machine and Deep Learning, Information and Network Security, Data mining & Data warehousing, NLP, and Measurement of Component-based Systems. He can be contacted at email: nasibsgill@gmail.com.