JURNAL TEKNOLOGI REKAYASA INFORMASI DAN KOMPUTER
E-ISSN: 2797-1724, DOI: 10.30811/jtrik.v9i1.8707
Vol. 9 No. 1 2026 pp. 25-33

Application of K-Means and Web-Based GIS For Poverty Mapping In The Aceh Region

Irwansyah[1], Mulyadi[2], Huzaeni[3]

[12,3] 3Department of Information Technology and Computer Science, Politeknik Negeri Lhokseumawe
Jln. B.Aceh Medan Km.280 Buketrata 24301 INDONESIA

*Penulis Korespondensi: irwansyahsyah271@gmail.com

Article info: Received on 09/10/2025, Revised 09/01/2026, Accepted 18/02/2026

Abstract

Poverty is a serious problem facing Indonesia, especially in developing regions such as Aceh, and can hinder national progress. Poverty involves various aspects, such as economic conditions, housing, and individual social capabilities. The variability of poverty data between regions is influenced by various factors, including social assistance, income, and other social factors. Therefore, solutions are needed to better manage and understand poverty data in Aceh. Two approaches that can be taken to overcome this challenge are the k-means clustering method and geographic information systems (GIS). The k-means clustering method is a statistical tool that allows researchers to group regions based on poverty levels, helping to identify relevant patterns. Meanwhile, GIS is used to analyse and visualise poverty data in the form of maps, facilitating understanding of poverty distribution patterns in Aceh. The development of a web-based geographic information system can also facilitate public and government access to poverty data in Aceh, increasing transparency and participation in addressing poverty issues. The results of this system will produce three groups of regions, namely non-poor, poor and very poor. Furthermore, this system has an accuracy of 93% similarity with the results conducted by BPS Lhokseumawe and Alu O Village.

Keywords: Poverty, Aceh, K-Means Clustering, Geographic Information System (GIS)

1. Introduction

Poverty is a very serious problem, especially in developing countries such as Indonesia. Poverty in a region can have a negative impact on national development, hindering economic progress and the overall welfare of society [1]. Poverty has many aspects that need to be considered. Economically, poverty is often identified as a lack of resources necessary to meet basic needs and improve quality of life. Housing is also a determining factor, with temporary or substandard housing often being an indicator of poverty. An individual's social and economic ability to obtain employment and meet their basic needs is also an important factor in measuring poverty [1].

Poverty data in each village or region can vary significantly. This is influenced by various factors, including the level of social assistance received, economic income, and other factors. Therefore, poverty alleviation in each region must be tailored to its unique conditions. In this context, grouping these regions based on established poverty characteristics is an important step. The 2020 poverty data for Aceh Province released by the Aceh Provincial Statistics Agency ( ) of the Central Statistics Agency (BPS) is used as a reference for this grouping [1].

One of the most striking problems in Aceh today is the lack of a system capable of organising poverty data properly and presenting it in various relevant groupings. So far, there has been no effective system for managing poverty data and presenting it to the public and government in a transparent and accessible manner. Therefore, a solution is needed to overcome this challenge and provide a better understanding of the poverty problem in Aceh [1].

To address this challenge, two approaches can be used: the K-means clustering method and geographic information systems (GIS). The K-means clustering method is a statistical tool that allows researchers to group regions based on their similar characteristics. With this method, regions can be grouped based on their poverty levels, helping to identify relevant patterns [1]. On the other hand, geographic information systems (GIS) can be used to analyse poverty levels in the region and visualise this information in the form of maps. Thus, GIS can assist the government and the community in understanding the patterns of poverty distribution in Aceh and taking appropriate measures to address this issue. By developing a web-based geographic information system that maps the distribution of poverty in Aceh, access to this information by the community and government can be improved, increasing transparency and participation in poverty alleviation efforts. This information can be accessed online via the internet [1].

By combining the K-means clustering method and geographic information systems, we can improve our understanding of poverty issues in Aceh and help the government and community take more effective measures to address this problem. In this journal, we will discuss the implementation of these two approaches to address poverty issues in Aceh in greater depth.

2. Method

2.1. System Overview

The system to be developed is a web-based Geographic Information System (GIS) that aims to map the distribution of poverty in Aceh. This system will use the K-means Clustering algorithm to group poverty data into relevant categories. Users will be able to view interactive maps that visualise the results of poverty data grouping based on various criteria, such as poverty levels, type of residence, or other social characteristics. The main objective of this system is to provide easy access to users from various levels of society and government to understand and take appropriate action in addressing poverty issues in the Aceh region.

2.2. Poverty

Poverty is the inability to meet basic needs, which can be caused by a scarcity of the means to meet those needs or difficulties in accessing education and employment. In East Aceh, the poverty rate reaches 14.08%, which is quite high for the entire Aceh region, especially in rural areas. Many villages in remote areas face poverty involving economic aspects, inadequate housing, and difficulty in finding employment [1].

2.3. Geographic Information System (GIS)

A Geographic Information System (GIS) is a system designed to collect, store, process, analyse, organise, and display various types of geographic data. The basic principles of GIS involve geographic information about locations on the earth's surface, the use of geographic information technology such as GPS, remote sensing, and specialised software, as well as various applications in different fields. Geographic Information Science is the basis for the development of GIS technology [2].

GIS is a computer system that can manage geographic data, including data collection, storage, retrieval, manipulation, analysis, and final results in the form of visualisation. GIS technology integrates database-based data processing operations with geographic analysis, providing advantages in data retrieval based on needs and unique geographic data visualisation [3].

GIS is a field of study and technology that continues to evolve, is used in various disciplines, and is experiencing rapid development. In general, GIS can be defined as an information system used to manage, analyse, and generate geographic data that supports decision-making in planning and management [3].

2.4. K-Means Clustering

The K-Means algorithm is a common and frequently used algorithm in industry. This algorithm begins by determining the number of clusters to be formed and selecting the initial centre point for each cluster. Then, the algorithm repeatedly performs the following steps until it reaches stability: first, determining the coordinates of the centre point of each cluster; second, measuring the distance of each object to the centre point coordinates; and third, grouping objects based on the closest distance. The results of this clustering process can be visualised as shown in Figure 1.

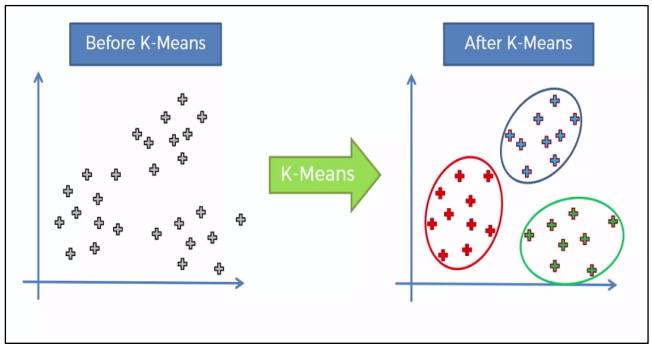

Figure 1. Illustration of Clustering Results

Figure 1 shows the clustering results using the K-Means method, where different colours represent different clusters. The main advantages of this method are its simplicity and speed, making it suitable for large datasets. However, this method has disadvantages, including uncertainty in the final results because it depends on random starting points and does not always produce the global minimum variant. In addition, this method requires a definable concept of average, which is not always fulfilled [9].

The clustering process using the K-Means Clustering method in this study involves several stages, as shown in Figure 2. The first stage is to determine the number of clusters (k), then generate random initial centroids. Next, distance calculations using the Euclidean formula are performed for each input data point relative to the centroid. The data is then grouped based on the closest distance, and the centroid value is updated based on the average of the cluster concerned. This process is repeated until there are no more changes in cluster members. The final result is the cluster centre value in the last iteration used as a parameter for data classification [8].

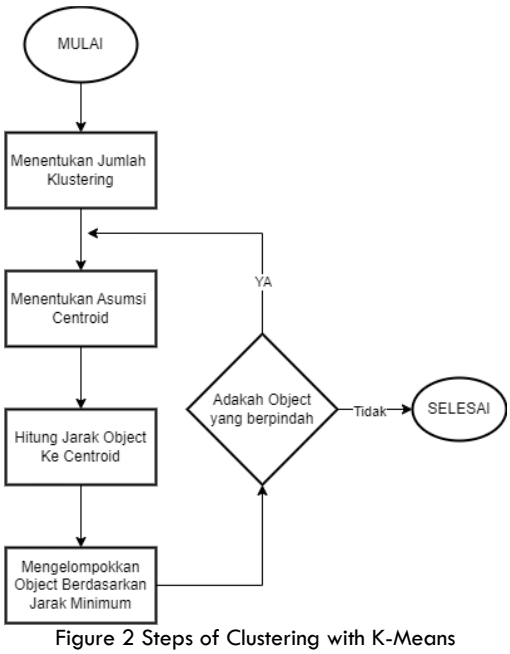

Figure 2 Steps of Clustering with K-Means

2.5. BlackBox Testing

Black Box Testing is a software testing method that focuses on testing based on requirement specifications without the need to examine or explore the source code. In this test, testers treat the software as a black box where they only have knowledge of the inputs and the expected outputs. This approach is carried out from the perspective of the customer or end user, so that the overall functionality of the system is evaluated. The main advantage of Black Box Testing is its ability to validate whether the software meets customer requirements, as well as detect any requirements that may be incomplete or unpredictable. In addition, this method allows for the testing of valid and invalid inputs to observe how the system handles various situations [10].

2.6. Data and data collection

The sample data used in this study was obtained from secondary data sources, namely the 2022 Poverty

Data in Aceh obtained from the Aceh Central Statistics Agency (BPS Aceh). Table 1 shows the details of the poverty level data used in this study.

Table 1 Sample poverty Data

No	Nama	Kabupaten	Kecamatan	Pendidikan	Pendapatan	Pekerjaan	Pangan
1	Umah	Simeulue	Alafan	SMP	1200000	Tidak bekerja	Sangat kurang
2	Saifullah	Aceh Singkil	Danau Paris	SD	100000	Pegawai BUMN	Melimpah
3	Abdullah Umar	Aceh Selatan	Kluet Timur	S2	1500000	Pegawai BUMD	Melimpah
4	Hanfiah	Aceh Tenggara	Babul Rahmah	D4	600000	Pegawai BUMN	Sangat melimpah
5	Hj. Maliana	Aceh Timur	Darul Falah	MIN	800000	Nelayan	Sangat melimpah
6	Samsul Bahri	Aceh Tengah	Bebesen	SMP	1300000	Nelayan	Sangat melimpah

2.7. UML

Unified Modelling Language (UML) is a visual language used in software engineering to design, document, and model software systems. UML is used to graphically describe various aspects of a software system, including system structure, interactions between components, system behaviour, and relationships between the entities involved.

1. Use Case Diagram

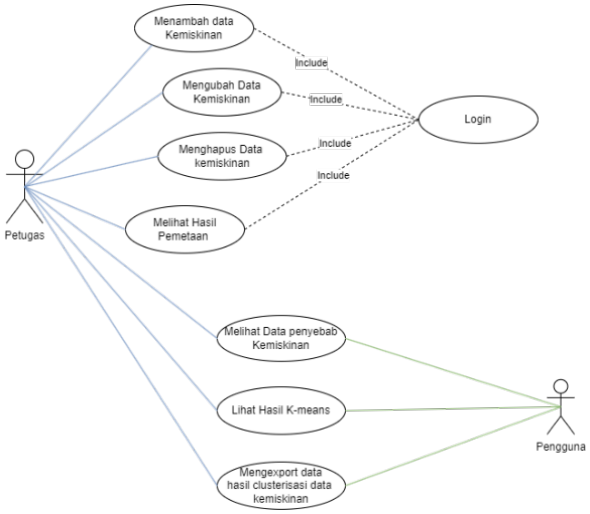

Figure 3. Use Case Diagram

Figure 3 illustrates the Use Case Diagram showing the interaction between "officials" and "users" in the system. There are ten Use Cases covering login, poverty data management by officials, viewing data and mapping results by officials and users, as well as other functions such as clustering and exporting clustered data. The Use Case Diagram provides a visual overview of the system's features and functions.

2. Activity Diagram
   a. Activity Diagram for GIS implementation : The process begins when the user selects the "Maps" option, whereupon the system responds by displaying a map page. After that, the user can select a specific area to view. Next, the system retrieves the relevant data from the database and displays the area selected by the user on the screen.

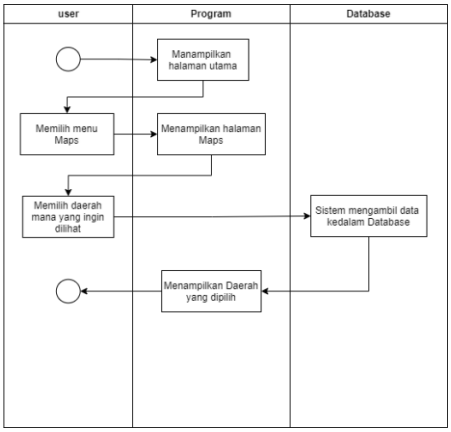

Figure 4 Activity Diagram of GIS Implementation

b. Poverty Data Export Diagram : This is an activity diagram for "Exporting poverty data clustering results." The process begins when the user selects the "Poverty Data" option from the menu. Next, the system will display poverty data on the screen. After that, both officers and users have the option to select the "Download" menu. Once the "Download" menu is selected, the system will initiate the process of generating data in PDF format and displaying the PDF to the user or officer.

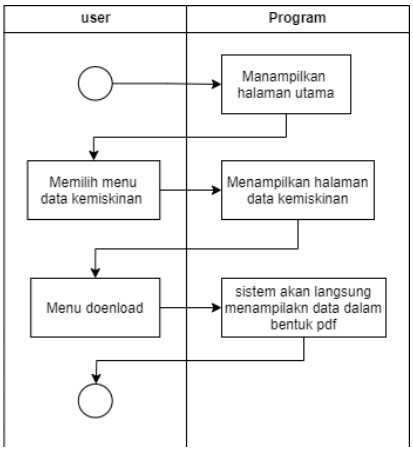

Figure 5 Activity Diagram for Poverty Data Export

c. Activity diagram for adding poverty data : The process begins when an officer or user selects the "Add Poverty Data" option from the menu. Next, the system will display the add poverty data page. The officer or user will enter the new data into the form provided. Once the data has been entered, the system will save the data to the database and display a success message to the user or officer to confirm that the data has been successfully saved.

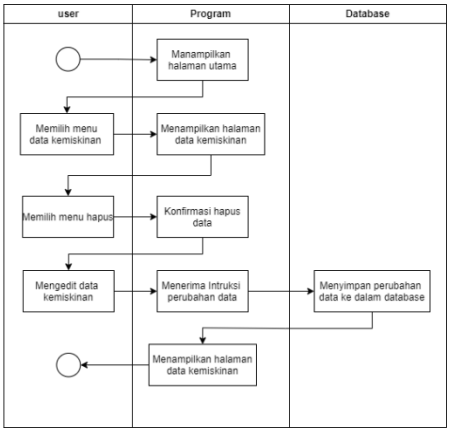

Figure 6 Activity diagram for adding poverty data

3. Results and Discussion

3.1. System Design Implementation

a. Implementation of the home page

The home page is the first screen that appears when users or officers access the system. In the image below, you can see the home page designed to provide users or officers with easy and intuitive access to various system features and functions

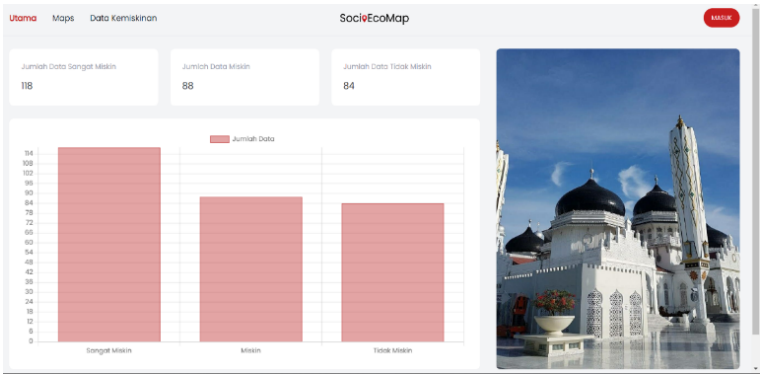

Figure 7. Implementation of the home page

b. Implementation of the poverty data page

In Figure 8, we can see the Poverty Data page, which is an important component of the system. On this page, poverty data has undergone a clustering process by the system to provide a deeper understanding of the patterns and groups of data. An interesting feature is the poverty data report that has been generated in PDF format. This report allows users to easily save, print, or share poverty data relevant to related parties. This, in turn, will assist in better decision-making and more effective strategic planning in addressing poverty issues. All these aspects create an informative and useful page display for poverty data management.

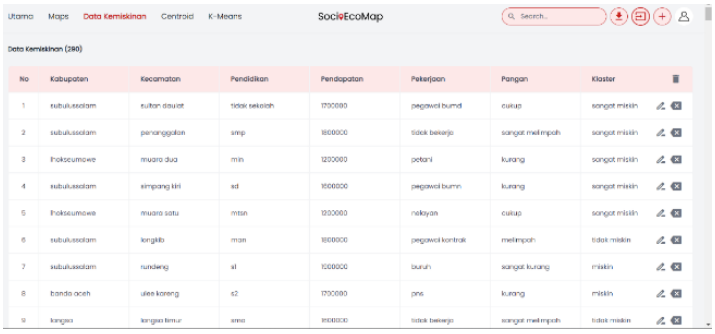

Figure 8 Application of the poverty data page

c. Implementasi GIS

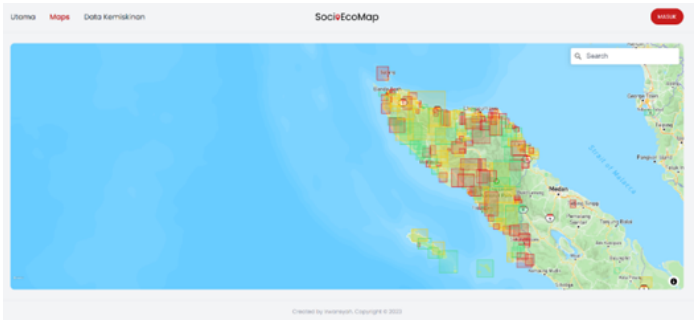

Figure 9. GIS

In Figure 4.2, we can see the Maps page displaying the province of Aceh along with 23 districts/cities, 290 sub-districts, and poverty data that has undergone a clustering process. For each sub-district, poverty data is displayed along with information about poverty category classifications, such as "poor", "very poor" and "not poor".

This provides a comprehensive view of the distribution of poverty in the region, allowing users to quickly identify and understand poverty patterns and levels in various locations.

3.2. Application of K-Means

The following are the steps for K-Means clustering in the web-based Poverty Mapping application:

1. The first step is to enter poverty data and relevant parameters, including Population Income, Food Availability, Employment Rate, and Education. This data will be used in the clustering process.
2. Next, we initialise or determine the number of clusters to be created. In this case, we start with three clusters. The number of clusters is determined based on analysis using the Elbow method and Silhouette score.
3. After that, we select data from two sub-districts as the starting point or initial centroid for each cluster. The parameters used to determine this centroid are Education, Income, Employment, and Food Availability. An example of a cluster centroid can be stated as follows: (example of centroid values)
4. The iteration process begins and continues until there are no changes in the placement of districts into clusters. The iteration involves the following steps:
    a. Calculating the Euclidean distance between each sub-district and the cluster centroid using the Euclidean distance formula.
    b. Calculating the distance between each sub-district and the two initial centroids.
    c. Determine which cluster has the closest Euclidean distance for each subdistrict.
    d. Update the centroid for each cluster based on the sub-district data included in the cluster's .
    e. If there are no changes in the placement of subdistricts into clusters in the next iteration, the convergence process has been achieved.
5. The final result of the clustering process is the grouping of districts into clusters based on poverty levels, namely Poor, Very Poor, and Non-Poor clusters.

3.3. Black Box Testing

Based on the black box testing in Appendix 1, the poverty mapping application produces the following data:

1. Total number of test cases: 12
2. All test cases passed and none failed.

To calculate the pass rate in black box testing, we can use the following formula:

Pk (%) = (Passing testing /Total Testing) x 100% (1)

By substituting the appropriate values into the formula, the pass rate for black box testing can be calculated as follows:

$$\text{Pk}\ (\%) = (12 \ / \ 12)\ x\ 100\% = 100\% \quad (2)$$

Thus, the black box test results for the mapping application have a pass rate of approximately 100%, and there are no failed test cases.

4. Conclusion

In this study, a web-based Geographic Information System (GIS) was successfully developed using the K-Means Clustering method to analyse the distribution of poverty in Aceh. The results show that GIS is effective in identifying patterns of poverty distribution. The web-based approach improves the accessibility of information for the government and the community, supporting decision-making related to poverty reduction. This study provides in-depth insights into the distribution of poverty in Aceh and has the potential to contribute to more efficient poverty eradication policies. Obstacles in managing geographic data must be overcome, and GIS technology and cluster analysis can be further utilised to understand the causes of poverty and develop more effective solutions.

REFERENSI

[1] M. Sigala, A. Beer, L. Hodgson, and A. O'Connor, *Big Data for Measuring the Impact of Tourism Economic Development Programmes: A Process and Quality Criteria Framework for Using Big Data*. 2019.

[2] G. Nguyen *et al.*, "Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey," *Artif. Intell. Rev.*, vol. 52, no. 1, pp. 77–124, 2019, doi: 10.1007/s10462-018-09679-z.

[3] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J. Big Data*, vol. 6, n[1] N. I. Febianto dan N. Palasara, "Analisa Clustering K-Means Pada Data Informasi Kemiskinan Di Jawa Barat Tahun 2018," Jurnal Sisfokom (Sistem Informasi dan Komputer), vol. 8, no. 2, hlm. 130–140, Agu

2019, doi: 10.32736/ sisfokom. v8i2.653.
[2] Y. Rahmanto dan S. Hotijah, "PERANCANGAN SISTEM INFORMASI GEOGRAFIS KEBUDAYAAN LAMPUNG BERBASIS MOBILE," 2020.
[3] O. M. Tumimomor, E. Jando, dan E. Meolbatak, "SISTEM INFORMASI GEOGRAFIS PARIWISATA KOTA KUPANG," Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI), vol. 1, no. 2, 2013.
[4] R. Puji Hastanti dan B. Eka Purnama, "Sistem Penjualan Berbasis Web (E-Commerce) Pada Tata Distro Kabupaten Pacitan," Jurnal Bianglala Informatika, vol. 3, no. 2, 2015, [Daring]. Tersedia pada: http://lppm3.bsi.ac.id/jurnal
[5] Yadi Utama, "SISTEM INFORMASI BERBASIS WEB JURUSAN SISTEM INFORMASI FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA," Jurnal Sistem Informasi (JSI), vol. 3, 2011, Diakses: 19 September 2023. [Daring]. Tersedia pada: http://ejournal.unsri.ac.id/index.php /jsi/index
[6] M. Silalahi dan D. Wahyudi, "Computer Based Information System Journal PERBANDINGAN PERFORMANSI DATABASE MONGODB DAN MYSQL DALAM APLIKASI FILE MULTIMEDIA BERBASIS WEB," CBIS JOURNAL, vol. 06, no. 01, 2018, [Daring]. Tersedia pada: http://ejournal.upbatam.ac.id/index.php/cbis\http://ejournal.upbatam.ac.id/index.php/cbis
[7] O. : Taryana dan S. M. Kom, "MENGENAL JAVA SCRIPT." [Daring]. Tersedia pada: https://us04web.zoom.us/j/ 3629329963?pwd=ZUhWcXI6RHp3dTNKZmlxWkUrV2ZCQT09
[8] S.Andayani, "Dipresentasikan dalam SEMNAS Matematika dan Pendidikan Matematika 2007 dengan tema 'Trend Penelitian Matematika dan Pendidikan Matematika di Era Global' yang diselenggarakan oleh Jurdik Matematika FMIPA UNY Yogyakarta pada tanggal 24 Nopember 2007."
[9] K. Singh, D. Malik, dan N. Sharma, "Evolving limitations in K-means algorithm in data mining and their removal," 2011. [Daring]. Tersedia pada: www.IJCEM.orgIJCEMwww.ijcem.org
[10] S. Nidhra, "Black Box and White Box Testing Techniques - A Literature Review," International Journal of Embedded Systems and Applications, vol. 2, no. 2, hlm. 29–50, Jun 2012, doi: 10.5121/ijesa.2012.2204.o. 1, 2019, doi: 10.1186/s40537-019-0197-0.
[4] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep Learning Approach for Intelligent Intrusion Detection System," *IEEE Access*, vol. 7, pp. 41525–41550, 2019, doi: 10.1109/ACCESS.2019.2895334.
[5] K. Sivaraman, R. M. V. Krishnan, B. Sundarraj, and S. Sri Gowthem, "Network failure detection and diagnosis by analyzing syslog and SNS data: Applying big data analysis to network operations," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 9 Special Issue 3, pp. 883–887, 2019, doi: 10.35940/ijitee.I3187.0789S319.
[6] A. D. Dwivedi, G. Srivastava, S. Dhar, and R. Singh, "A decentralized privacy-preserving healthcare blockchain for IoT," *Sensors (Switzerland)*, vol. 19, no. 2, pp. 1–17, 2019, doi: 10.3390/s19020326.
[7] F. Al-Turjman, H. Zahmatkesh, and L. Mostarda, "Quantifying uncertainty in internet of medical things and big-data services using intelligence and deep learning," *IEEE Access*, vol. 7, pp. 115749–115759, 2019, doi: 10.1109/ACCESS.2019.2931637.
[8] S. Kumar and M. Singh, "Big data analytics for healthcare industry: Impact, applications, and tools," *Big Data Min. Anal.*, vol. 2, no. 1, pp. 48–57, 2019, doi: 10.26599/BDMA.2018.9020031.
[9] L. M. Ang, K. P. Seng, G. K. Ijemaru, and A. M. Zungeru, "Deployment of IoV for Smart Cities: Applications, Architecture, and Challenges," *IEEE Access*, vol. 7, pp. 6473–6492, 2019, doi: 10.1109/ACCESS.2018.2887076.
[10] B. P. L. Lau *et al.*, "A survey of data fusion in smart city applications," *Inf. Fusion*, vol. 52, no. January, pp. 357–374, 2019, doi: 10.1016/j.inffus.2019.05.004.
[11] Y. Wu *et al.*, "Large scale incremental learning," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2019-June, pp. 374–382, 2019, doi: 10.1109/CVPR.2019.00046.
[12] A. Mosavi, S. Shamshirband, E. Salwana, K. wing Chau, and J. H. M. Tah, "Prediction of multi-inputs bubble column reactor using a novel hybrid model of computational fluid dynamics and machine learning," *Eng. Appl. Comput. Fluid Mech.*, vol. 13, no. 1, pp. 482–492, 2019, doi: 10.1080/19942060.2019.1613448.
[13] V. Palanisamy and R. Thirunavukarasu, "Implications of big data analytics in developing healthcare frameworks – A review," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 31, no. 4, pp. 415–425, 2019, doi: 10.1016/j.jksuci.2017.12.007.
[14] J. Sadowski, "When data is capital: Datafication, accumulation, and extraction," *Big Data Soc.*, vol. 6, no. 1, pp. 1–12, 2019, doi: 10.1177/2053951718820549.
[15] J. R. Saura, B. R. Herraez, and A. Reyes-Menendez, "Comparing a traditional approach for financial brand communication analysis with a big data analytics technique," *IEEE Access*, vol. 7, pp. 37100–37108, 2019, doi: 10.1109/ACCESS.2019.2905301.
[16] D. Nallaperuma *et al.*, "Online Incremental Machine Learning Platform for Big Data-Driven Smart Traffic Management," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 12, pp. 4679–4690, 2019, doi: 10.1109/TITS.2019.2924883.
[17] S. Schulz, M. Becker, M. R. Groseclose, S. Schadt, and C. Hopf, "Advanced MALDI mass spectrometry imaging in pharmaceutical research and drug development," *Curr. Opin. Biotechnol.*, vol. 55, pp. 51–59, 2019, doi:

10.1016/j.copbio.2018.08.003.
[18] C. Shang and F. You, "Data Analytics and Machine Learning for Smart Process Manufacturing: Recent Advances and Perspectives in the Big Data Era," *Engineering*, vol. 5, no. 6, pp. 1010–1016, 2019, doi: 10.1016/j.eng.2019.01.019.
[19] Y. Yu, M. Li, L. Liu, Y. Li, and J. Wang, "Clinical big data and deep learning: Applications, challenges, and future outlooks," *Big Data Min. Anal.*, vol. 2, no. 4, pp. 288–305, 2019, doi: 10.26599/BDMA.2019.9020007.
[20] M. Huang, W. Liu, T. Wang, H. Song, X. Li, and A. Liu, "A queuing delay utilization scheme for on-path service aggregation in services-oriented computing networks," *IEEE Access*, vol. 7, pp. 23816–23833, 2019, doi: 10.1109/ACCESS.2019.2899402.
[21] G. Xu, Y. Shi, X. Sun, and W. Shen, "Internet of things in marine environment monitoring: A review," *Sensors (Switzerland)*, vol. 19, no. 7, pp. 1–21, 2019, doi: 10.3390/s19071711.
[22] M. Aqib, R. Mehmood, A. Alzahrani, I. Katib, A. Albeshri, and S. M. Altowaijri, *Smarter traffic prediction using big data, in-memory computing, deep learning and gpus*, vol. 19, no. 9. 2019.
[23] S. Leonelli and N. Tempini, *Data Journeys in the Sciences*. 2020.
[24] N. Stylos and J. Zwiegelaar, *Big Data as a Game Changer: How Does It Shape Business Intelligence Within a Tourism and Hospitality Industry Context?* 2019.
[25] Q. Song, H. Ge, J. Caverlee, and X. Hu, "Tensor completion algorithms in big data analytics," *arXiv*, vol. 13, no. 1, 2017.