Application of the outlier detection method for web-based blood glucose level monitoring system

Rachma Aurya Nurhaliza[1], Muhammad Qois Huzyan Octava[1], Farhan Mufti Hilmy[1], Umar Farooq[2], Ganjar Alfian[1]
[1]Department of Electrical Engineering and Informatics, Vocational College, Universitas Gadjah Mada, Yogyakarta, Indonesia
[2]Faculty of Business and Law, Coventry University, Coventry, United Kingdom

Article Info

Article history:

Received Oct 15, 2023
Revised Feb 15, 2024
Accepted Feb 23, 2024

Keywords:

3-sigma
Diabetes
Interquartile range
Monitoring system
Outlier detection

ABSTRACT

Recent advancements in biosensors have empowered individuals with diabetes to autonomously monitor their blood glucose levels through continuous glucose monitoring (CGM) sensors. Nevertheless, the data collected from these sensors may occasionally include outliers due to the inherent imperfections of the sensor devices. Consequently, the identification of these outliers is critical to determine whether blood glucose levels deviate significantly from the norm, necessitating further action. This study employs an outlier detection approach based on the 3-sigma method and the interquartile range (IQR), along with the application of the Winsorizing technique to correct the identified outliers. Additionally, a web-based system for visualizing blood glucose levels is developed, utilizing both outlier detection methods. In order to assess the system's performance, two types of testing are conducted: black box testing and load testing. The results of black box testing indicate that all test scenarios operate as anticipated. As for the load testing response times, it is observed that the 3-sigma visualization page loads an average of 606.75 milliseconds faster compared to the IQR visualization page. This study's outcomes are expected to enhance data quality, enhance the precision of analyses, and facilitate more informed decision-making by identifying and addressing extreme data points.

Corresponding Author:

Ganjar Alfian
Department of Electrical Engineering and Informatics, Vocational College, Universitas Gadjah Mada
Yogyakarta 55281, Indonesia
Email: ganjar.alfian@ugm.ac.id

1. INTRODUCTION

Diabetes is a chronic condition characterized by abnormal blood glucose metabolism. In type 1 diabetes, the pancreas loses its ability to produce insulin, necessitating insulin injections for patients to regulate their blood glucose levels. In contrast, type 2 diabetes involves the body's cells becoming less responsive to insulin, even though the body can still produce it [1]-[3]. Therefore, type 2 diabetes patients are advised to manage their diet, engage in physical exercise, and may be prescribed medication to control blood glucose levels. It is crucial to maintain blood glucose levels within a specific range, ensuring they do not fall below 70 mg/dL, which is known as hypoglycemia, or exceed 180 mg/dL, referred to as hyperglycemia [4]. Hyperglycemia can lead to complications such as kidney disease and cardiovascular issues [5], [6], making diabetics susceptible to stroke and a higher mortality rate [7]. Continuous monitoring of blood glucose levels can play a pivotal role in reducing and preventing complications in diabetic patients [8].

In 2012, diabetes was accountable for 1.5 million fatalities, along with an additional 2.2 million deaths attributed to conditions like cardiovascular disease, stroke, and other illnesses arising from complications related to diabetes [9]. To mitigate and prevent these complications associated with diabetes, modern information technology, along with biosensors, now enables the continuous monitoring of blood glucose level changes using portable self-monitoring of blood glucose (SMBG) devices and real-time continuous glucose monitoring (CGM) sensors [10]. This real-time monitoring facilitates immediate responses from healthcare professionals in case any issues arise [11]. The CGM sensor is a technology designed for evaluating blood glucose levels, utilizing an internet of things (IoT)-based approach to provide real-time measurements of blood glucose levels [12].

Data from the CGM device's sensors may contain outliers, which are data points that significantly deviate from the norm or, in simpler terms, do not follow the expected data pattern [13]. The presence of outliers in the data can introduce bias into statistical calculations, such as the mean, resulting in values that are either too low or too high [14]. Outlier detection is the process of identifying unusual patterns in the data that do not conform to the expected standard data. These outlier detection methods can be employed during data preprocessing stage to identify data irregularities or items considered as outliers [15], [16]. In the context of sensor data, outliers arise due to imperfections in the sensing devices and weaknesses in network connectivity, which can lead to some of the collected data being incomplete [17].

Numerous studies in the field of outlier detection have been carried out by various researchers. One such investigation was conducted by Tallon-Ballesteros and Riquelme [18], who employed the interquartile range (IQR) outlier detection method to assess the impact of outliers in classification tasks and shed light on the contentious matter of outlier removal. Their study demonstrated that the application of the IQR outlier detection method substantially enhanced the performance of the C4.5 classification algorithm when some outliers were selectively removed from the training dataset. In contrast, Capelleveen *et al.* [19] compared linear models, IQR, and multivariate clustering techniques to identify outliers in real-world data within the Medicaid dental insurance domain. Their research concluded that IQR is more suitable for identifying fraudulent patterns in healthcare systems. Another comparative study on outlier detection methods was conducted by Jeong *et al.* [20] where they compared the 3-sigma, IQR, and mean absolute deviation (MAD) approaches in the context of groundwater analysis. Their findings indicated that the IQR method outperformed both 3-sigma and MAD in terms of performance. Furthermore, Nascimento *et al.* [21] evaluates various outlier detection techniques, such as boxplot and 3-sigma, for identifying outliers in power consumption data from a tertiary building in France, concluding that combining a regression method like random forest with the adjusted boxplot method holds promising potential for detecting this type of data quality issue in electricity consumption.

Nonetheless, there has been no previous investigation into evaluating 3-sigma and IQR outlier detection methods for monitoring blood glucose levels from CGM device. Therefore, this study utilizes statistical techniques to detect outliers in the data from blood glucose CGM sensors and employs the Winsorizing technique to rectify such outliers. Detecting outliers is vital as it is essential to rectify these outliers in the dataset to enhance the predictive model's performance. Furthermore, a web-based system for visualizing patients' blood glucose level data will be developed, with the capability to deliver accurate information to users. The system will present blood glucose data through a line chart visualization. Additionally, the performance of the web-based system will be evaluated through software testing, encompassing both black box testing and load testing.

2. METHOD

The process analysis comprises multiple phases, commencing with a review of existing literature, gathering pertinent data, performing data preprocessing through the utilization of outlier detection methods and Winsorizing techniques, formulating the system's design, integrating it into the system, and conducting system testing. As can be seen in Figure 1, the initial phases of this study commenced with a comprehensive review of literature sources, encompassing research journals, books, and online resources, to gather relevant information. Following the acquisition of the CGM dataset, the subsequent step involves data preprocessing. In this study, the data will undergo processing using the 3-sigma and IQR methods. Once any outliers are identified in the data, the outlier data will undergo further processing using the Winsorizing technique. The subsequent phase entails system design, which includes creating use cases to capture the dynamic behavior of the system and designing the user interface (UI) to delineate how the system will be developed. The implementation phase involves constructing the system according to the established design, using widely used web development PHP programming language and managing data with MySQL as the database [22]. After the web-based system becomes operational, a series of tests are conducted to ascertain whether the system operates as expected. In this study, two types of system testing were conducted, namely, black box testing and load testing.

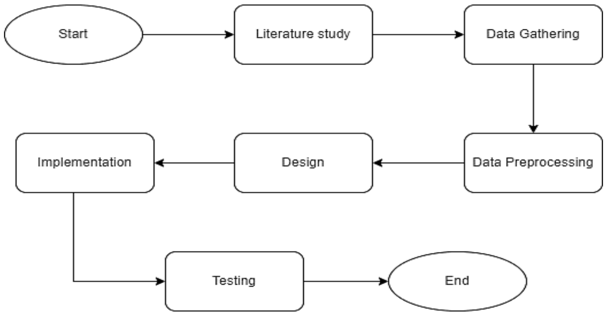


Figure 1. Analysis process

2.1. Dataset

In this study, the CGM dataset consisted of time series data derived from the records of blood glucose levels in 30 pediatric patients diagnosed with type 1 diabetes. All participants, children diagnosed with type 1 diabetes, utilized the Guardian-RT device, a CGM system, to collect blood glucose data at 5-minute intervals over a period of about 7 days. This dataset was sourced from a study titled "Evaluation of Counterregulatory Hormone Responses during Hypoglycemia and the Accuracy of Continuous Glucose Monitors in Children with T1DM," which was conducted by the Jaeb Center for Health Research (JCHR) [23]. The dataset encompasses 56,719 data entries and includes fields for record ID, patient ID, date and time of recording, blood glucose level, calibration status, and file type.

2.2. Pre-processing stage

2.2.1. 3-sigma method

The 3-sigma method is a statistical technique that involves data falling within three standard deviations of the mean. It is employed to establish upper and lower boundaries in statistical computations. Data points exceeding the mean plus three times the standard deviation, or falling below the mean minus three times the standard deviation, are identified as outliers, while data points within these upper and lower limits are deemed normal [24]. The values for these upper and lower limit lines in the 3-sigma method can be determined using (1) and (2). Given $\bar{x}$ is the average value (mean) of all data, σ is the standard deviation of all data, x_i is the value in the i-th data, and *n* is the amount of data, the calculation of mean and standard deviation is obtained by (3) and (4).

$$\text{Upper Limit}=\bar{x}+3\sigma \tag{1}$$

$$\text{Lower Limit}=\bar{x}-3\sigma \tag{2}$$

$$\bar{x}=\frac{\sum x_i}{n} \tag{3}$$

$$\sigma=\frac{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}}{n-1} \tag{4}$$

2.2.2. Interquartile range method

The IQR method is a statistical approach employed to gauge the dispersion of a dataset and to identify outliers. The IQR corresponds to the difference between the lower quartile (25th percentile) and the upper quartile (75th percentile) of the data. Initially, the data is organized, and the quartile values are computed. Subsequently, the IQR is calculated by subtracting the value of the lower quartile from that of the upper quartile. To detect outliers, the lower limit is established by subtracting 1.5 times the IQR from the lower quartile, while the upper limit is determined by adding 1.5 times the IQR to the upper quartile. Data points outside of these limits are classified as outliers [25]. The values for these upper and lower limit lines in the IQR method can be determined using (5) and (6). Given *n* as the quantity of data, the calculation of Q1, Q3, and IQR values are obtained by (7) to (9).

$$\text{Upper limit}=Q3+(1{,}5\times IQR) \tag{5}$$

$$\text{Lower limit}=Q1-(1{,}5\times IQR) \tag{6}$$

$$\text{Position of } Q1=\frac{1}{4}(n+1) \tag{7}$$

$$\text{Position of Q3}=\frac{3}{4}(n+1) \quad (8)$$

$$IQR=Q3–Q1 \quad (9)$$

2.2.3. Winsorizing

Winsorizing is a statistical procedure involving the substitution of outlier values within the dataset to mitigate the potential impact of outliers. The benefit of Winsorizing lies in its ability to replace outlier values with either the highest or lowest values within the data distribution while safeguarding the data against some of the adverse effects of outliers [25], [26]. In this study, the Winsorizing technique is applied to address outlier data by setting it equal to either the upper limit or lower limit value from 3-sigma or IQR method. In doing so, the Winsorizing technique enhances the robustness of statistical analyses by reducing the influence of outliers while preserving the overall integrity of the dataset.

3. RESULTS AND DISCUSSION

The CGM dataset will undergo processing with two distinct outlier detection techniques, namely, the 3-sigma and IQR methods. Each of these methods will compute upper and lower limit values. Subsequently, any data points within the dataset exceeding these upper or lower limits, as determined by both methods, will be identified as outliers. These outlier data points will then be adjusted based on the upper or lower limit values set by each method, employing the Winsorizing technique. Moreover, the web-based system will display the data that has been altered through Winsorizing as a line chart visualization.

3.1. Comparison of outlier detection method outcomes

Following the application of the 3-sigma and IQR methods to the sensor data, variations emerge in the outcomes of outlier detection pertaining to patient blood glucose levels. As depicted in the Table 1, it is evident that the IQR-based outlier detection method identifies a higher number of outliers compared to the 3-sigma method.

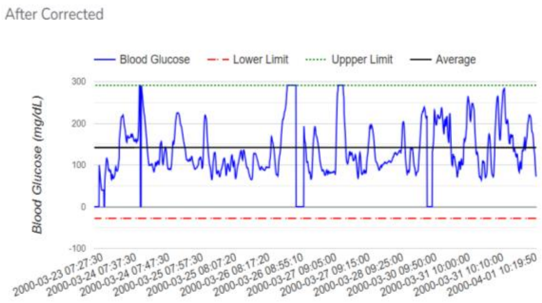
Table 1. Comparison of outlier detection method outcomes

No	Patient ID	Number of outliers detected by 3-sigma method	Number of outliers detected by IQR method
1	1	184	8
2	3	0	0
3	4	0	0
4	5	0	0
5	6	0	0
6	8	7	14
7	9	0	1
8	10	21	36
9	11	164	164
10	12	44	63
11	13	0	0
12	14	0	115
13	15	152	152
14	16	7	6
15	17	0	89
16	18	35	45
17	19	0	0
18	20	0	0
19	21	52	76
20	22	109	116
21	23	9	15
22	24	64	65
23	25	183	0
24	26	0	0
25	28	0	0
26	29	4	24
27	30	0	0
28	31	0	0
29	32	0	13
30	33	11	50

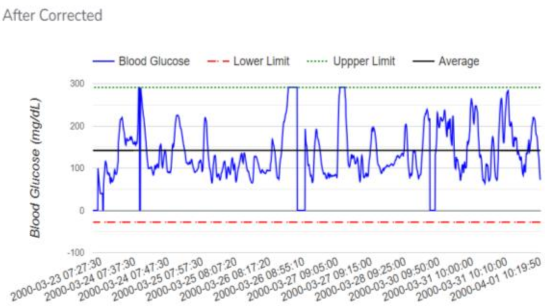
Furthermore, this disparity in outlier detection highlights the significance of selecting an appropriate method based on the specific requirements and objectives of the analysis. The choice between these methods should be made carefully to ensure accurate identification of outliers in patient blood glucose level data. In practical terms, this implies that the IQR method may be more sensitive to deviations in blood glucose levels, which could be valuable in clinical settings to ensure patient safety. Researchers and healthcare professionals should consider the trade-off between sensitivity and specificity when selecting an outlier detection method, as an increased number of detections may also introduce some false positives.

3.2. Practical application

The ultimate aim of this study is to integrate the 3-sigma and IQR outlier detection methods within a web-based system designed for visualizing the correct value of blood glucose. This system will serve the purpose of alerting users, primarily administrators, to the presence of outliers or irregular blood glucose readings in patients. Users will have the ability to view corrected blood glucose level visualizations based on patient IDs. The system will utilize either the 3-sigma or IQR method formulas in conjunction with the Winsorizing technique, all accessible through the web application interface. Figure 2(a) presents a system page demonstrating the visual representation of blood glucose levels and the application of the 3-sigma method as well as Winsorizing technique for patient 12. Conversely, Figure 2(b) displays a similar page illustrating blood glucose visualization after employing IQR and Winsorizing for patient 12. These pages exhibit visualizations in the form of line charts, complete with data on blood glucose levels, date and time of recording, upper and lower limits, as well as the average values.



(a)



(b)

Figure 2. Practical application for; (a) visualization of blood glucose data using 3-sigma method and (b) using IQR method

3.3. System testing

System testing is a critical phase in software development where the entire software system is evaluated as a whole to ensure that all its components work together as intended. The primary benefit of system testing is the assurance that the software will perform reliably in real-world conditions, providing users with a robust and error-free experience. This study involved two categories of testing: functional and non-functional. Functional testing will utilize black box testing, while non-functional testing will employ load testing.

3.3.1. Black box testing

Black box testing is a software testing approach focused on evaluating the system's functionality through its user interface. This method examines the software based on specifications, without delving into the program's code or internal structure. Its primary purpose is to assess how the system responds to specific inputs and identify any issues or defects in the software [27]. Table 2 showed the test scenarios and its result of black box testing.

Table 2. Black box testing scenario

Test scenario	Test case	Expected results	Actual results	Status
View the outlier visualization for patient 12 in the 3-sigma visualization menu	Click the 3-sigma Visualization menu and select patient 12	The system display the outlier visualization in the form of a line chart of patient 12	The system displays the outlier visualization in the form of a line chart of patient 12	Success
View the outlier visualization for patient 12 in the IQR visualization menu	Click the IQR Visualization menu and select patient 12	The system display the outlier visualization in the form of a line chart of patient 12	The system displays the outlier visualization in the form of a line chart of patient 12	Success
View the outlier data for patient 12 within the 3-sigma outlier data menu	Click the 3-sigma outlier data menu and select patient 12	The system display the patient 12 blood glucose level outlier data table	The system displays the patient 12 blood glucose level outlier data table	Success
View the outlier data for patient 12 within the IQR outlier data menu	Click the IQR outlier data menu and select patient 12	The system display the patient 12 blood glucose level outlier data table	The system displays the patient 12 blood glucose level outlier data table	Success

The outcome indicated that in black box testing, all test scenarios were executed as anticipated. This positive result in black box testing suggests that the software's specified functionalities are working correctly, aligning with the intended user experience. It provides confidence that the system's user interface and overall behavior meet the defined requirements.

3.3.2. Load testing

Load testing, as a type of non-functional testing, is a performance evaluation method used to assess how a system responds to diverse user requests. It aids in understanding how the software performs when multiple users access the system concurrently, revealing potential bottlenecks or issues related to scalability [28]. Load testing is conducted on an actual system, whether it is in the form of a prototype or a fully operational system, rather than a design on paper [29]. In this particular study, load testing was conducted using the JMeter software, which simulates server requests by creating numerous simultaneous user HTTP request threads. The key performance metrics evaluated by JMeter include response time, throughput, latency, response bytes, and load time [30]. Figure 3 presents a visualization of the comparative results obtained from load testing, showcasing the performance outcomes for various numbers of users.

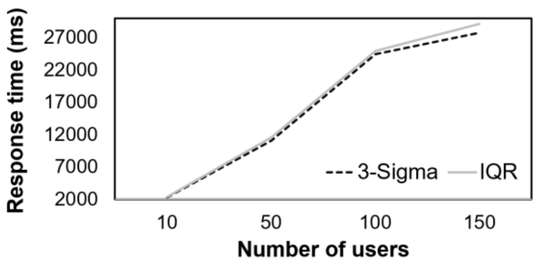


Figure 3. Visualization of load testing comparison outcomes

We have performed experiments with various user numbers accessing the system, including 10, 50, 100, and 150 users. The results indicated that the 3-sigma visualization page consistently outperformed the IQR visualization page, with an average response time 606.75 milliseconds faster. Additionally, an increase in user load leads to an associated increase in response times. These findings highlight the importance of considering the choice of visualization method, as the 3-sigma page appears to offer a more responsive user experience. It also emphasizes the need for efficient resource management to maintain optimal performance as user demand grows.

4. CONCLUSION

The development of a web-based system for visualizing blood glucose levels, using PHP and a MySQL database, was successful. The system effectively identifies outlier data through both the 3-sigma and IQR methods, with the IQR method detecting more outliers. The identified outlier data is then processed using the Winsorizing technique. The results, including the visualization displayed in line charts, are integrated into the web-based system. This web-based blood glucose visualization system underwent two types of testing: black box testing, which ran all test scenarios as expected, and load testing. The results showed that during black box testing, all the test scenarios involving the integration of IQR and 3-sigma into the web application were carried out as expected. During load testing, the 3-sigma visualization page consistently demonstrated superior performance compared to the IQR visualization page, exhibiting an average response time that was 606.75 milliseconds faster. Nevertheless, the system's visualization remained quick and responsive for both outlier detection methods, ensuring overall performance was not compromised. Future research will involve the utilization of different datasets and the integration of outlier detection into prediction models, with the aim of enhancing prediction performance.

REFERENCES

[1] "2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2018," *Diabetes Care*, vol. 41, no. Supplement_1, pp. S13–S27, Jan. 2018, doi: 10.2337/dc18-S002.

[2] J. Silverstein *et al.*, "Care of Children and Adolescents With Type 1 Diabetes," *Diabetes Care*, vol. 28, no. 1, pp. 186–212, Jan. 2005, doi: 10.2337/diacare.28.1.186.

[3] N. A. ElSayed *et al.*, "2. Classification and Diagnosis of Diabetes: Standards of Care in Diabetes—2023," *Diabetes Care*, vol. 46, no. Supplement_1, pp. S19–S40, Jan. 2023, doi: 10.2337/dc23-S002.

[4] D. E. Goldstein *et al.*, "Tests of Glycemia in Diabetes," *Diabetes Care*, vol. 27, no. 7, pp. 1761–1773, Jul. 2004, doi: 10.2337/diacare.27.7.1761.

[5] R. Z. Alicic, M. T. Rooney, and K. R. Tuttle, "Diabetic Kidney Disease," *Clinical Journal of the American Society of Nephrology*, vol. 12, no. 12, pp. 2032–2045, Dec. 2017, doi: 10.2215/CJN.11491116.

[6] V. Ormazabal, S. Nair, O. Elfeky, C. Aguayo, C. Salomon, and F. A. Zuñiga, "Association between insulin resistance and the development of cardiovascular disease," *Cardiovascular Diabetology*, vol. 17, no. 1, pp. 1–14, Dec. 2018, doi: 10.1186/s12933-018-0762-4.

[7] N. N. Tun, G. Arunagirinathan, S. K. Munshi, and J. M. Pappachan, "Diabetes mellitus and stroke: A clinical update," *World Journal of Diabetes*, vol. 8, no. 6, p. 235, 2017, doi: 10.4239/wjd.v8.i6.235.

[8] S. H. Ley, O. Hamdy, V. Mohan, and F. B. Hu, "Prevention and management of type 2 diabetes: dietary components and nutritional strategies," *The Lancet*, vol. 383, no. 9933, pp. 1999–2007, Jun. 2014, doi: 10.1016/S0140-6736(14)60613-9.

[9] "Global report on diabetes," *World Health Organization*. 2016. [Online]. Available: https://iris.who.int/handle/10665/204871. (accessed: Feb. 12, 2024).

[10] J. Al-Nabulsi *et al.*, "Non-invasive sensing techniques for glucose detection: a review," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 4, pp. 1926–1937, Aug. 2022, doi: 10.11591/eei.v11i4.3584.

[11] D. Bruen, C. Delaney, L. Florea, and D. Diamond, "Glucose Sensing for Diabetes Monitoring: Recent Developments," *Sensors*, vol. 17, no. 8, p. 1866, Aug. 2017, doi: 10.3390/s17081866.

[12] D. Olczuk and R. Priefer, "A history of continuous glucose monitors (CGMs) in self-monitoring of diabetes mellitus," *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 12, no. 2, pp. 181–187, Apr. 2018, doi: 10.1016/j.dsx.2017.09.005.

[13] V. Barnett and T. Lewis, *Outliers in Statistical Data*. Singapore: John Wiley & Sons, 1994, doi: 10.2307/2344655.

[14] S. K. Kwak and J. H. Kim, "Statistical data preparation: management of missing values and outliers," *Korean Journal of Anesthesiology*, vol. 70, no. 4, p. 407, 2017, doi: 10.4097/kjae.2017.70.4.407.

[15] M. Ijaz, G. Alfian, M. Syafrudin, and J. Rhee, "Hybrid Prediction Model for Type 2 Diabetes and Hypertension Using DBSCAN-Based Outlier Detection, Synthetic Minority Over Sampling Technique (SMOTE), and Random Forest," *Applied Sciences*, vol. 8, no. 8, p. 1325, Aug. 2018, doi: 10.3390/app8081325.

[16] M. T. Alasaady, T. N. M. Aris, N. M. Sharef, and H. Hamdan, "A proposed approach for diabetes diagnosis using neuro-fuzzy technique," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 6, pp. 3590–3597, Dec. 2022, doi: 10.11591/eei.v11i6.4269.

[17] J. W. Branch, C. Giannella, B. Szymanski, R. Wolff, and H. Kargupta, "In-network outlier detection in wireless sensor networks," *Knowledge and Information Systems*, vol. 34, no. 1, pp. 23–54, Jan. 2013, doi: 10.1007/s10115-011-0474-5.

[18] A. J. Tallon-Ballesteros and J. C. Riquelme, "Deleting or keeping outliers for classifier training?," in *2014 Sixth World Congress on Nature and Biologically Inspired Computing (NaBIC 2014)*, Jul. 2014, pp. 281–286, doi: 10.1109/NaBIC.2014.6921892.

[19] G. van Capelleveen, M. Poel, R. M. Mueller, D. Thornton, and J. van Hillegersberg, "Outlier detection in healthcare fraud: A case study in the Medicaid dental domain," *International Journal of Accounting Information Systems*, vol. 21, pp. 18–31, Jun. 2016, doi: 10.1016/j.accinf.2016.04.001.

[20] J. Jeong, E. Park, W. S. Han, K. Kim, S. Choung, and I. M. Chung, "Identifying outliers of non-Gaussian groundwater state data based on ensemble estimation for long-term trends," *Journal of Hydrology*, vol. 548, pp. 135–144, May 2017, doi: 10.1016/j.jhydrol.2017.02.058.

[21] G. F. M. Nascimento, F. Wurtz, P. Kuo-Peng, B. Delinchant, and N. J. Batistela, "Outlier Detection in Buildings' Power Consumption Data Using Forecast Error," *Energies*, vol. 14, no. 24, p. 8325, Dec. 2021, doi: 10.3390/en14248325.

[22] L. Welling and L. Thomson, *PHP and MySQL Web development*. Sams publishing, 2005.

[23] "Diabetes Research in Children Network (DirecNet)," *JCHR*. [Online]. Available: https://public.jaeb.org/direcnet/stdy/167. (accessed: Oct. 01, 2022).

[24] R. Lehmann, "3σ-Rule for Outlier Detection from the Viewpoint of Geodetic Adjustment," *Journal of Surveying Engineering*, vol. 139, no. 4, pp. 157–165, Nov. 2013, doi: 10.1061/(ASCE)SU.1943-5428.0000112.

[25] C. S. K. Dash, A. K. Behera, S. Dehuri, and A. Ghosh, "An outliers detection and elimination framework in classification task of data mining," *Decision Analytics Journal*, vol. 6, p. 100164, Mar. 2023, doi: 10.1016/j.dajour.2023.100164.

[26] J. C. Adams, D. K. Hayunga, S. Mansi, and D. M. Reeb, "Identifying Outliers in Finance," *SSRN Electronic Journal*, vol. 48, no. 2, pp. 345–384, 2019, doi: 10.2139/ssrn.2986928.

[27] B. Beizer and J. Wiley, "Black Box Testing: Techniques for Functional Testing of Software and Systems," *IEEE Software*, vol. 13, no. 5, p. 98, Sep. 1996, doi: 10.1109/MS.1996.536464.

[28] B. L. Romano *et al.*, "Software Testing for Web-Applications Non-Functional Requirements," in *2009 Sixth International Conference on Information Technology: New Generations*, 2009, pp. 1674–1675, doi: 10.1109/ITNG.2009.209.

[29] Z. M. Jiang and A. E. Hassan, "A Survey on Load Testing of Large-Scale Software Systems," *IEEE Transactions on Software Engineering*, vol. 41, no. 11, pp. 1091–1118, Nov. 2015, doi: 10.1109/TSE.2015.2445340.

[30] V. Tiwari, S. Upadhyay, J. K. Goswami, and S. Agrawal, "Analytical Evaluation of Web Performance Testing Tools: Apache JMeter and SoapUI," in *2023 IEEE 12th International Conference on Communication Systems and Network Technologies (CSNT)*, Apr. 2023, pp. 519–523, doi: 10.1109/CSNT57126.2023.10134699.

BIOGRAPHIES OF AUTHORS

Rachma Aurya Nurhaliza currently is a fresh graduate from Department of Electrical Engineering and Informatics, Vocational College, Universitas Gadjah Mada, majoring in Software Engineering Technology. Her research areas are artificial intelligence, machine learning, and web development. She can be contacted at email: rachmaaurya@mail.ugm.ac.id.

Muhammad Qois Huzyan Octava currently is a fresh graduate from Department of Electrical Engineering and Informatics, Vocational College, Universitas Gadjah Mada, majoring in Software Engineering Technology. His research areas are artificial intelligence, data mining, and software development. He can be contacted at email: qoisoctava@mail.ugm.ac.id.

Farhan Mufti Hilmy currently is a fresh graduate from Department of Electrical Engineering and Informatics, Vocational College, Universitas Gadjah Mada, majoring in Software Engineering Technology. His research areas are artificial intelligence, machine learning, and customer relationship management. He can be contacted at email: farhanmuftihilmy@gmail.com.

Umar Farooq is currently working as a lecturer of Business Analytics at the School of Strategy and Leadership, College of Business and Law, Coventry University, United Kingdom. Previously, he worked as Assistant Professor for five year at Gulam Ishaq Khan Institute of Engineering Sciences and Technology, Pakistan. He received his Masters and Ph.D. in Industrial and Systems Engineering from Dongguk University, South Korea in 2017. He has vast experience in the fields of supply chain management, operations management, and business analytics. His research interests are in RFID, IoT, food supply chains, machine learning, and analytics. He can be contacted at email: ad9061@coventry.ac.uk.

Ganjar Alfian currently works as Assistant Professor at Department of Electrical Engineering and Informatics, Vocational College, Universitas Gadjah Mada, Indonesia since 2022. Previously, he worked as Assistant Professor at Dongguk University Seoul, Republic of Korea for five years. In July 2017, he was a short-term Visiting Researcher with the VSB-Technical University of Ostrava, Czech Republic. He received Dr.Eng. degrees from the Department of Industrial and Systems Engineering, Dongguk University, Seoul, South Korea, in 2016. His research interests include artificial intelligence, RFID, IoT, machine learning, carsharing, and simulation. He is listed in the World's Top 2% Scientist 2022 by Stanford University. He can be contacted at email: ganjar.alfian@ugm.ac.id.