

Improving Internet of Things Platform with Anomaly Detection for Environmental Sensor Data

Okyza Maherdy Prabowo¹, Suhono Harso Supangkat², Eueung Mulyana³, I Gusti Bagus Baskara Nugraha⁴

STMIK AMIK Bandung, Bandung Institute of Technology, Bandung, Indonesia¹

Smart City and Community Innovation Center, Bandung Institute of Technology, Bandung, Indonesia²

School of Electrical Engineering and Informatics, Bandung Institute of Technology, Bandung, Indonesia^{3,4}

Abstract—Internet of things has an essential role in various application domains. The number of Internet of Things applications makes researchers try to formulate how to design the architecture of the Internet of Things platform so that it can be used generically in various domains. Commonly used architectural designs consist of data collecting, data preprocessing, data analysis, and data visualization. However, sensor data that enters the platform often experiences anomalies such as constant values or being stuck-at zero, which are processed manually at the data preprocessing stage. In this research, we try to design an anomaly detection system on the Internet of Things platform that can automatically improve the platform's performance in detecting anomalies. In this study, we compared the False Positive Rate of several anomaly detection algorithms tested to real datasets in the environmental sensor data application domain. The results showed that the anomaly detector system on the Internet of Things platform had an optimal False Positive Rate of 0.9%.

Keywords—Anomaly detection; sensor data; multivariate; Internet of Things; smart system

I. INTRODUCTION

Recently some urban environments have extensively used internet of things (IoT) technology to perform environmental monitoring and control. The acquisition and control settings and the network protocols vary according to the urban environment's intended applications. These elements are critical to the ability of IoT networks to communicate successfully and transfer valuable data. Valuable data such as air temperature and relative humidity from inside and outdoor locations are essential for understanding the urban microclimate affecting the environmental condition. The monitoring process needs help from technology tools to automate the collection and understanding of data, for example, the internet of things platform.

The platform collects microclimate parameters from all sensor data. The platform also serves as a data management platform. The data platform architecture has a subsystem called the data analytics module. The data analytics module is responsible for analyzing the collected data. The data analytics module can be implemented through video, text, or other analytics techniques such as statistical analysis or machine learning [1]. Anomaly detection is one of the analyses performed on the data collected in an agricultural environment [2].

The common goal of anomaly detection is to find patterns in data that do not conform to “expected” or normal behavior [3]. Anomaly detection is used to monitor the environmental situation of the greenhouse [19]. When anomalous behavior is detected, an alarm can be sent to the administrator to do something. Several techniques are used to detect anomalies, which can be classified into two categories: conventional and data-driven. A conventional technique like the statistical method has a long history of detecting an outlier in the data [20]. Parametric or non-parametric techniques are included in this category. The underlying distribution of the data is known for the parametric category, and the parameters are estimated using the data. Parametric methods include those based on the Gaussian distribution, the regression model, or a combination of Parametric Distributions [4]. Data-driven techniques are frequently used to refer to learning-based methods in which the lack of a robust underlying mathematical model is compensated for by the availability of large amounts of data from which useful information can be “learned.” Machine learning is a large area of research with numerous application areas. Generally, it is divided into three distinct categories: supervised, unsupervised, and reinforcement learning. Additionally, due to technological advancements, deep learning is gaining popularity. Numerous machine learning techniques are frequently given a deep learning orientation or are combined with deep learning [18].

Several algorithms for detecting sensor abnormalities are used in agricultural environments. Several neural network algorithms were used, including artificial neural networks, autoencoders, recurrent neural networks, and long short-term memory. However, in complex environments where a clear variation pattern for some greenhouse parameters is complicated, environmental anomalies are rarely captured or recognized by univariate sensor data analysis or single machine learning models [5]. A multivariate anomaly detection approach is needed to be explored in Internet of Things area [16]. The anomaly detector system proposed in this study uses a GRU-based Variational Autoencoder. Guo proposed this method to handle IoT sensor data in Smart City [7]. The advantage of the GRU-based anomaly detection system is its reliability in discovering the data correlation and dependencies [14]. There are still weaknesses in actual labelling, which will be improved in this study by involving human knowledge as part of a multivariate anomaly detection system. The main contribution in this paper are summarized as follows:

- Anomaly detector framework for the Internet of Things platform handling type of sensors error combining data-driven and knowledge-driven.
- Algorithm comparison results that fit the Internet of things platform performing real-world datasets in the greenhouse system in the tropical country.

The rest of this article is organized as follows. The Gated Recurrent Unit (GRU) and Variational Autoencoder (VAE)-based anomaly detector are described in Section II, along with the proposed architecture combining GRU-VAE and Human-in-the-loop method. By incorporating knowledge into data-driven techniques, we can increase the detection rate of interesting anomalies [11]. Section III contains the results and discussion for three datasets. Finally, Section IV contains some concluding remarks.

II. SYSTEM MODEL

In this section, the Gated Recurrent Unit and Variational Autoencoder-based anomaly detection are given along with the proposed model.

A. Gated Recurrent Unit

The GRU, or gated recurrent unit, is an improvement over the RNN, or recurrent neural network. It was first introduced by Cho et al. in 2014[6]. GRUs are strikingly similar to Long-Short-Term Memory (LSTM). Similar to LSTM, GRU utilizes gates to control the information flow. In comparison to LSTM, they are a more recent development. Consequently, they outperform LSTM and have a more straightforward architecture, as referenced in Fig. 1.

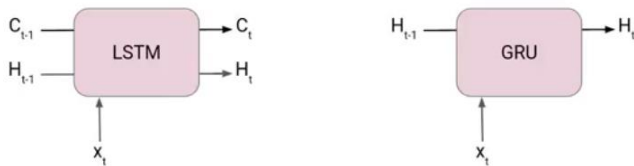


Fig. 1. Illustration Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) Cell.

At each timestamp t , it accepts the input X_t and the hidden state H_{t-1} from the previous timestamp $t-1$. The new hidden state H_t is then output and passed to the subsequent timestamp. As shown in Figures 1 and 2, a GRU cell consists primarily of two gates instead of three in an LSTM cell. The reset gate comes first, followed by the update gate. The reset gate is responsible for the network's short-term memory, known as its hidden state (H_t). The equation for the reset gate is as follows.

$$r_t = \sigma(x_t * U_r + H_{t-1} * W_r) \quad (1)$$

The update gate for long-term memory is illustrated below, along with the gate's equation.

$$u_t = \sigma(x_t * U_u + H_{t-1} * W_u) \quad (2)$$

The only distinction is between the weight metrics U_u and W_u . To locate the hidden state H_t in GRU, a two-step procedure is used. The first step is to create a candidate's hidden state. The hidden gate formula is described.

$$H_t = \tanh(x_t * U_g + (r_t \circ H_{t-1}) * W_u) \quad (3)$$

It multiplies the input and hidden state from the previous timestamp $t-1$ by the output of the reset gate r_t . This information is then passed to the tanh function, which returns the hidden state of the candidate. Important in this equation is how we use the reset gate's value to limit the previous hidden state's influence on the candidate state. If r_t equals 1, the previous hidden state H_{t-1} is evaluated in its entirety. Similarly, if r_t is 0, the information from the previous hidden state is completely disregarded. After determining the candidate state, it is used to generate the current hidden state H_t . The Update gate enters the fray at this point. This equation is highly intriguing because, unlike LSTM, we control the historical information in H_{t-1} and the new information in the candidate state with a single update gate.

$$H_t = u_t \circ H_{t-1} + (1 - u_t) * H_t \quad (4)$$

Now, if U_t is close to 0, the first term in the equation will vanish, implying that the new hidden state will contain little information about the previous hidden state. On the other hand, the second part becomes nearly identical to the first, which implies that the hidden state at the current timestamp will contain only information from the candidate state. Guo uses GRU cells in both the encoder and the decoder to discover the data correlation and dependency [7].

B. Variational Autoencoder based Anomaly Detection

Anomaly detection is one of those domains where machine learning has had such a profound impact that it is almost axiomatic that anomaly detection systems must be based on some type of automatic pattern learning algorithm as opposed to a set of rules or descriptive statistics (though many reliable anomaly detection systems operate using such methods very successfully and efficiently). Combining Bayesian inference with an AE framework, VAE is a probabilistic model. As opposed to a reconstruction error, a VAE-based anomaly detection model generates a probabilistic measure for the anomaly score [14]. Reconstruction probabilities are more objective and principled than reconstruction errors because they do not require modeling specific thresholds for judging anomalies. In particular, VAE assumes that a large number of complex data distributions can be described by a smaller set of latent variables with more straightforward probability density distributions. Thus, the objective of VAE is to find a low-dimensional representation of the latent variables in the input data.

The VAE is distinct from conventional autoencoders because it is probabilistic and generative. The VAE generates partially random outputs (even after training) and can also generate new data similar to the data on which it was trained. The VAE is structurally similar to a conventional autoencoder at a high level. However, the encoder acquires additional coding; specifically, the VAE acquires mean and standard deviation coding. The VAE then generates the latent variables, z , by randomly sampling from a Gaussian distribution with the same mean and standard deviation as the encoder. To reconstruct the input, these latent variables are "decoded." The architecture is visualized by Fig. 2.

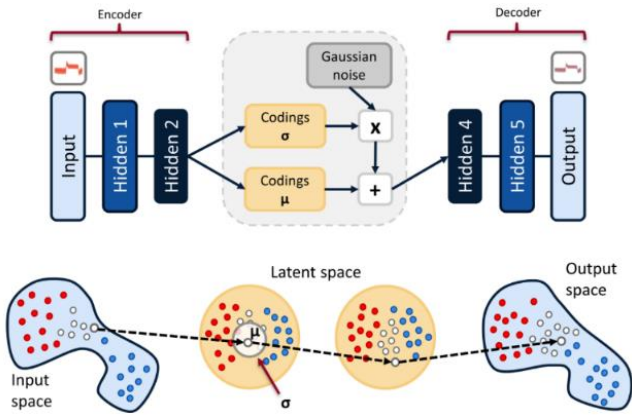


Fig. 2. A Variational Autoencoder Architecture (Top), and an Example of a Data Sample Going through the VAE (Bottom).

C. Human Knowledge-Driven

Human-driven machine intelligence is also known as "human in the loop." Human-in-the-loop (HITL) is a subfield of artificial intelligence that combines human and machine intelligence to create machine learning models [12]. In a conventional human-in-the-loop approach mentioned in Fig. 3, people are involved in a virtuous circle in which they train, tune, and test a specific algorithm. Humans initially assign labels to data, which provides a model with high-quality (and massive) training data. A machine learning algorithm learns to make decisions based on this data. Afterward, humans fine-tune the model. Humans typically score data to account for overfitting, to teach a classifier about edge cases, or to add new categories to the model's scope. Individuals can evaluate and validate a model's outputs, especially when an algorithm is uncertain about a judgment or overconfident about an incorrect decision [10].

D. Proposed Architecture

The proposed architecture for anomaly detection combines data-driven methods, specifically a GRU-based Variational Autoencoder, with expert-provided knowledge, as referenced in Fig. 4. The GRU-based variational autoencoder performs the anomaly detection process on the provided dataset and then compares it to the expert's knowledge [13]. On the greenhouse dataset, this knowledge-driven approach is used. The following diagram illustrates a multivariate anomaly detector's general architecture. The GRU-based VAE algorithm detects anomalies from multivariate sensors. Similarly, experts interpret anomalies in multivariate data.

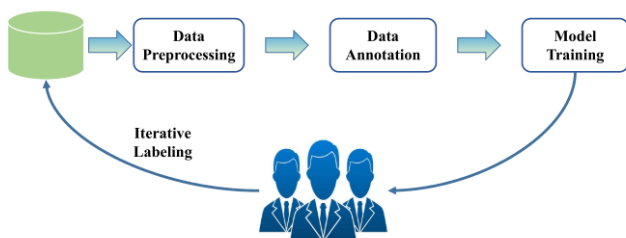


Fig. 3. The Development Cycle of Model [10].

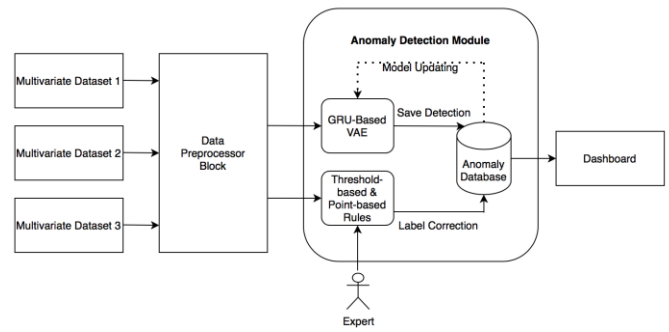


Fig. 4. Fusion Architecture Combining Data-driven and Knowledge-Driven.

This multivariate dataset is derived from three datasets: Intel Berkeley Dataset, indoor greenhouse sensors, and outdoor greenhouse sensors. Greenhouse sensors are being installed in Bandung, Indonesia, in the tropics. The data is cleaned of empty values in the preprocessing subsystem, and each dataset is resampled every 20 minutes. Following that, the feature scaling process is carried out, which is rescaling features to make them more suitable for training [15].

Data Preprocessing blocks are used to carry out several data management processes such as joining data, removing missing values, separating data according to sensor categories, resampling time and then entering into two different system blocks, namely data-driven block and human knowledge-driven block[17].

A data-driven anomaly detection architecture will be proposed in this study, which will make use of a GRU-based variational autoencoder with multivariate time-series data as input. First, greenhouse data is loaded from the database and split into two dataset categories, indoor module, and outdoor module. Each sensor contains four sensor variables, a battery sensor, a humidity sensor, and two temperature sensors.

The GRU input accepts four sensor inputs, each of which is connected to 150 cells in the first layer. Then, the output of the second layer is connected to the second layer's 100 inputs.

The human-in-the-loop method is used for knowledge-driven anomaly detection. Experts make label recommendations based on data using threshold-based and point-based methods. This threshold-based approach will be used to improve the anomaly detection threshold generated by the GRU-based VAE in the future. At the same time, the point-based is used by iteratively providing data to the expert and then labeling the points. The point-based anomaly detection process is measured by the amount of time it takes the expert to label it versus the amount of anomaly it takes to help the expert label it.

III. EXPERIMENT RESULT AND DISCUSSION

In this section, the proposed GRU-based VAE model is evaluated using the Intel Berkeley, greenhouse indoor sensor, and greenhouse outdoor sensor datasets. Shamshiri proposed microclimate parameters to be evaluated [8]. Four criteria, namely accuracy, the area under curve (AUC), true positive rate, and false positive rate, are used to evaluate the performance. All experiments were run on the Google Colab

with Intel Xeon @ 2.2GHz, 12 GB Ram, and 12GB NVIDIA Tesla K80 GPU. The algorithm was implemented using Python in Keras and Scikit-learn. The expert knowledge-driven method is evaluated using greenhouse indoor sensor and greenhouse outdoor sensor dataset. The expert will be given a set of preprocessed data and labeled them. The time execution and the number of anomaly data will be compared as the evaluation metric. The expert knowledge-driven method performs better when the time required to determine the number of known anomalies is reduced.

A. Intel Berkeley Dataset

This dataset was compiled using data from 54 sensors installed in the Intel Berkeley Research lab between February 28th and April 5th, 2004 [9]. Every 31 seconds, it contains time stamped topology information, humidity, temperature, light, and voltage values. For various sensors, there are some missing values at certain timestamps. To begin, we use the linear interpretation method to fill in the gaps. Then, every 20 minutes, we sample it and use the average as input. To balance the type of sensor that fits the greenhouse sensor, we only take temperature, humidity, light, and voltage sensors. In the meantime, we normalize the data.

With GRU-based VAE, the average testing MAE is 0.04 and MSE is 0.01 with training time 194s. Table I shows GRU-Based VAE performance. The testing accuracy, Area Under Curve (AUC), Optimal True Positive Rate (Opt. TPR) and False Positive Rate (Opt. FPR) of 81%, 69%, 4%, and 1.8% for light sensor, 39%, 89%, 14.9% and 0% for voltage sensor, 34%, 84%, 5% and 1.6% for humidity sensor and 25%, 82%, 8.4% and 1.4% for temperature sensor, respectively.

TABLE I. GRU-BASED VAE PERFORMANCE

GRU-Based VAE				
Sensor Type	Accuracy	AUC	Opt. TPR	Opt. FPR
Light Sensor	81%	69%	4%	1.8%
Voltage Sensor	39%	89%	14.9%	0%
Humidity Sensor	34%	84%	5%	1.6%
Temperature Sensor	25%	82%	8.4%	1.4%

With Gaussian Mixture Model, the testing accuracy, Area Under Curve, Optimal True Positive Rate (Opt. TPR) and False Positive Rate (Opt. FPR) of 61%, 65%, 3.7%, and 1.2% for light sensor, 27%, 94%, 12.8% and 0% for voltage sensor, 32%, 84%, 4.9%, and 1.7% for humidity sensor and 23%, 49%, 0% and 23.7% for temperature sensor, respectively. Table II summarizes the evaluation result.

TABLE II. GAUSSIAN MIXTURE MODEL PERFORMANCE

Gaussian Mixture Model				
Sensor Type	Accuracy	AUC	Opt. TPR	Opt. FPR
Light Sensor	70%	65%	3%	1.8%
Voltage Sensor	73%	7%	0%	1%
Humidity Sensor	68%	16%	2%	4.9%
Temperature Sensor	25%	51%	2.4%	0%

With K-Means, the testing accuracy, Area Under Curve, Optimal True Positive Rate (Opt. TPR) and False Positive Rate (Opt. FPR) of 70%, 65%, 3.3%, and 1.8% for light sensor, 73%, 7%, 0% and 1.29% for voltage sensor, 68%, 16%, 1.7%, and 4.9% for humidity sensor and 78%, 51%, 2.4% and 0% for temperature sensor, respectively. Table III summarizes the evaluation result.

TABLE III. K-MEANS PERFORMANCE

K-Means				
Sensor Type	Accuracy	AUC	Opt. TPR	Opt. FPR
Light Sensor	70%	65%	3%	1.8%
Voltage Sensor	73%	7%	0%	1.3%
Humidity Sensor	68%	16%	1.7%	4.9%
Temperature Sensor	78%	51%	2.4%	0%

Based on the multivariate correlation between all sensors, there are no concurrent anomalies on the four sensors. There are 0.625% concurrent anomalies on the three sensors. There are 8.125% concurrent anomalies on the two sensors.

B. Greenhouse Indoor Dataset

This dataset was compiled using data from four sensors installed outside the Greenhouse Smart City Living Lab between October 16th, 2020, and July 19th, 2021. It contains timestamped information about the topology every 60 seconds, as well as humidity, two temperatures with distinct locations, and voltage values. There are some values missing at certain timestamps for various sensors. To begin, we will fill in the gaps using the linear interpretation technique. After that, it is sampled every 20 minutes, and the average is used as an input. We take two temperatures with a different locations, humidity, and voltage sensors to balance the sensor type that fits the greenhouse sensor. Meanwhile, we standardize the data.

With GRU-based VAE, the average testing MAE is 0.015 and MSE is 0.16 with training time 196s. Table IV shows GRU-Based VAE performance. The testing accuracy, Area Under Curve (AUC), Optimal True Positive Rate (Opt. TPR) and False Positive Rate (Opt. FPR) of 82%, 58%, 7%, and 3.2% for battery sensor, 93%, 74%, 8% and 1.4% for humidity sensor, 70%, 68%, 2.1% and 0.9% for temperature sensor DS-type and 71%, 70%, 6.4% and 2.8% for temperature sensor SHT-type, respectively.

TABLE IV. GRU-BASED VAE PERFORMANCE

GRU-Based VAE				
Sensor Type	Accuracy	AUC	Opt. TPR	Opt. FPR
Battery Sensor	82%	58%	7%	3.2%
Humidity Sensor	93%	74%	8%	1%
Temperature Sensor DS	70%	68%	2.1%	0.9%
Temperature Sensor SHT	71%	70%	6.4%	2.8%

With Gaussian Mixture Model, the testing accuracy, Area Under Curve, Optimal True Positive Rate (Opt. TPR) and False Positive Rate (Opt. FPR) of 64%, 55%, 5.5%, and 2.9% for battery sensor, 98%, NaN, 0% and 1.8% for humidity sensor,

66%, 69%, 0.2%, and 3.4% for temperature sensor DS-type and 69%, 72%, 1.4% and 8.8% for temperature sensor SHT-type, respectively. Table V summarizes the evaluation result.

TABLE V. GAUSSIAN MIXTURE MODEL PERFORMANCE

Gaussian Mixture Model				
Sensor Type	Accuracy	AUC	Opt. TPR	Opt. FPR
Battery Sensor	64%	55%	5.5%	2.9%
Humidity Sensor	98%	NaN	0%	2%
Temperature Sensor DS	66%	69%	0.2%	3.4%
Temperature Sensor SHT	69%	72%	1.4%	8.8%

With K-Means, the testing accuracy, Area Under Curve, Optimal True Positive Rate (Opt. TPR) and False Positive Rate (Opt. FPR) of 51%, 54%, 4.9%, and 2.7% for battery sensor, 98%, 100%, 100% and 1.8% for humidity sensor, 68%, 69%, 3.3%, and 2% for temperature sensor DS-type and 70%, 71%, 8.7% and 1.6% for temperature sensor SHT-type, respectively. Table VI summarizes the evaluation result.

TABLE VI. K-MEANS PERFORMANCE

K-Means				
Sensor Type	Accuracy	AUC	Opt. TPR	Opt. FPR
Battery Sensor	51%	54%	4.9%	2.7%
Humidity Sensor	98%	100%	100%	1.8%
Temperature Sensor DS	68%	69%	3.3%	2%
Temperature Sensor SHT	70%	71%	8.7%	1.6%

Based on the multivariate correlation between all sensors, there are 0.024% concurrent anomalies on the four sensors. There are 0.16% concurrent anomalies on the three sensors. There are 1.85% concurrent anomalies on the two sensors.

C. Greenhouse Outdoor Dataset

This dataset was compiled using data from 4 sensors installed in the Greenhouse Smart City Living Lab between October 16th, 2020, and July 19th, 2021. It also contains timestamped information about the topology every 60 seconds, as well as humidity, two temperatures with distinct locations, and voltage values. There are some values missing at certain timestamps for various sensors. To begin, we will fill in the gaps using the linear interpretation technique. After that, it is sampled every 20 minutes, and the average is used as an input. We take two temperatures with a different locations, humidity, and voltage sensors to balance the sensor type that fits the greenhouse sensor. Meanwhile, we standardize the data.

With GRU-based VAE, the average testing MAE is 0.05 and MSE is 0.007 with training time 72s. Table I shows GRU-Based VAE performance. The testing accuracy, Area Under Curve (AUC), Optimal True Positive Rate (Opt. TPR) and False Positive Rate (Opt. FPR) of 86%, 59%, 5.3%, and 1.1% for battery sensor, 78%, 60%, 1.5% and 1.5% for humidity sensor, 68%, 69%, 4.7% and 1% for temperature sensor DS-type and 69%, 63%, 1.2% and 0.6% for temperature sensor SHT-type, respectively.

TABLE VII. GRU-BASED VAE PERFORMANCE

GRU-Based VAE				
Sensor Type	Accuracy	AUC	Opt. TPR	Opt. FPR
Battery Sensor	86%	59%	5.3%	1.1%
Humidity Sensor	78%	60%	1.5%	1.5%
Temperature Sensor DS	68%	69%	4.7%	1%
Temperature Sensor SHT	69%	63%	1.2%	0.6%

With Gaussian Mixture Model, the testing accuracy, Area Under Curve, Optimal True Positive Rate (Opt. TPR) and False Positive Rate (Opt. FPR) of 61%, 57%, 3.3%, and 0.7% for battery sensor, 29%, 65%, 1.7% and 0.9% for humidity sensor, 68%, 67%, 5.4%, and 0.6% for temperature sensor DS-type and 73%, 62%, 1% and 0.7% for temperature sensor SHT-type, respectively. Table VIII summarizes the evaluation result.

TABLE VIII. GAUSSIAN MIXTURE MODEL PERFORMANCE

Gaussian Mixture Model				
Sensor Type	Accuracy	AUC	Opt. TPR	Opt. FPR
Battery Sensor	61%	57%	3.3%	0.7%
Humidity Sensor	29%	65%	1.7%	0.9%
Temperature Sensor DS	68%	67%	5.4%	0.6%
Temperature Sensor SHT	73%	62%	1%	0.7%

With K-Means, the testing accuracy, Area Under Curve, Optimal True Positive Rate (Opt. TPR) and False Positive Rate (Opt. FPR) of 48%, 57%, 0.7%, and 2.6% for battery sensor, 73%, 62%, 1.9% and 1.3% for humidity sensor, 68%, 67%, 5.5%, and 0.6% for temperature sensor DS-type and 72%, 63%, 0.7% and 0.9% for temperature sensor SHT-type, respectively. Table IX summarizes the evaluation result.

TABLE IX. K-MEANS PERFORMANCE

K-Means				
Sensor Type	Accuracy	AUC	Opt. TPR	Opt. FPR
Battery Sensor	48%	57%	0.7%	2.6%
Humidity Sensor	73%	62%	1.9%	1.3%
Temperature Sensor DS	68%	67%	5.5%	0.6%
Temperature Sensor SHT	72%	63%	0.7%	0.9%

Based on the multivariate correlation between all sensors, there are 0.13% concurrent anomalies on the four sensors. There are 0.19% concurrent anomalies on the three sensors. There are 0.58% concurrent anomalies on the two sensors.

D. Threshold-based & Point-based Human Knowledge Driven

This paper introduces point-based anomaly detection as part of a proposed method for determining how human-in-the-loop evaluation can be performed. This proposed method describes how an expert can provide anomaly recommendations through the threshold and point annotations. An agricultural expert was involved in determining the point anomaly in the Greenhouse Smart City Living Lab context in this greenhouse case study.

Experts are given raw data in stages, and the amount of raw data given is used to determine how much raw data is required to make it easier for experts to annotate anomalies. The time the data is displayed before the expert can provide annotations is then calculated. Based on the results of tests with data ranging from $n = 1$ to $n = 100$, it was discovered that the optimal expert produced the fastest results with $n = 5$ and an annotation time of 7.12 seconds. That is, the expert requires a minimum of five data samples in order to draw an annotation conclusion. Human knowledge is used as an adaptive threshold in the threshold-based approach, which can replace sigma, which is currently used as a threshold limit. With human knowledge stored in the database, the anomaly detection process will become more adaptive by adjusting the context or rules provided by humans based on point annotations or specific conditions such as a crop disease [20].

E. Discussion

GRU-based VAE has performed well in detecting anomalies, particularly the relationship between the detected variables. However, GRU-based VAE does not produce the best results in some datasets because it necessitates layer adjustments based on the data conditions. However, the deficiency in the anomaly detection process is compensated for by the assistance of human knowledge. Unlike the other algorithms, it has not been able to demonstrate the correlation between anomalies from multiple sensors simultaneously. However, a more detailed assessment of this correlation is required. Correlation is only indicated in this study by the classifications of no correlation or correlates.

This study also proposes a new metric for measuring human-in-the-loop by comparing the amount of data required for annotation and the time it takes the expert to annotate. The comparison curve of the anomaly n and the required time t is generally close to a quadratic function, making it difficult for the expert to annotate due to a lack of data. However, having too much data will also make it difficult for the expert.

IV. CONCLUSION

The greenhouse, outfitted with sensors, generates a large amount of data that must be processed. Of course, the data cannot be separated from anomalies, which may be an anomaly that must be removed because it corrupts the data, or the anomaly may represent hidden information that can be used to make future decisions. The Gated Recurrent Unit-based Variational Autoencoder is proposed in this study as an anomaly detection algorithm capable of detecting anomalies in the multivariate term. This algorithm is also a component of the anomaly detection architecture, which is enhanced by threshold-based and point-based anomaly detection based on human knowledge, which can improve anomaly detection performance.

This anomaly detection model and architecture were evaluated using the Intel Berkeley Lab Dataset, Greenhouse Smart City Living Lab Dataset, and indoor and outdoor sensors. The evaluation results demonstrate that the proposed model is superior at detecting multivariate anomalies and identifying variable correlation. Our proposed architecture using GRU-based VAE and expert feedback can examine

correlations between multivariable time series data. The human knowledge module enhances the performance of the GRU-based VAE by correcting false alarms and detecting errors.

Future works are necessary to validate the kinds of conclusions that can be drawn from this research. For example, it is necessary to measure point-based and threshold data for future case studies that may generate different curves. In addition, it is anticipated that this human knowledge will be utilized automatically to enhance the GRU-based VAE anomaly detection model in the future research.

REFERENCES

- [1] Erhan, L., Ndubuaku, M., Di Mauro, M., Song, W., Chen, M., Fortino, G., Bagdasar, O., & Liotta, A. (2021). Smart anomaly detection in sensor systems: A multi-perspective review. *Information Fusion*, 67, 64–79. <https://doi.org/10.1016/j.inffus.2020.10.001>.
- [2] Ou, C. H., Chen, Y. A., Huang, T. W., & Huang, N. F. (2020). Design and Implementation of Anomaly Condition Detection in Agricultural IoT Platform System. *International Conference on Information Networking*, 2020-Januari, 184–189. <https://doi.org/10.1109/ICOIN48656.2020.9016618>.
- [3] Liu, Y., Pang, Z., Karlsson, M., & Gong, S. (2020). Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control. *Building and Environment*, 183. <https://doi.org/10.1016/j.buildenv.2020.107212>.
- [4] Farzad, A., & Gulliver, T. A. (2020). Unsupervised log message anomaly detection. *ICT Express*, 6(3), 229–237. <https://doi.org/10.1016/j.ict.2020.06.003>.
- [5] Liu, Y., Pang, Z., Karlsson, M., & Gong, S. (2020). Anomaly detection based on machine learning in IoT-based vertical plant wall for indoor climate control. *Building and Environment*, 183, 107212. <https://linkinghub.elsevier.com/retrieve/pii/S0360132320305837>.
- [6] Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. 1–9. <http://arxiv.org/abs/1412.3555>.
- [7] Guo, Y., Ji, T., Wang, Q., Yu, L., Min, G., & Li, P. (2020). Unsupervised Anomaly Detection in IoT Systems for Smart Cities. *IEEE Transactions on Network Science and Engineering*, 7(4), 2231–2242. <https://doi.org/10.1109/TNSE.2020.302754>.
- [8] Shamshiri, R. R., Bojic, I., van Henten, E., Balasundram, S. K., Dworak, V., Sultan, M., & Weltzien, C. (2020). Model-based evaluation of greenhouse microclimate using IoT-Sensor data fusion for energy efficient crop production. *Journal of Cleaner Production*, 263, 121303. <https://doi.org/10.1016/j.jclepro.2020.121303>.
- [9] Kingma, D. P., & Welling, M. (2014). Auto-encoding variational bayes. 2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings, MI, 1–14.
- [10] Wu, X., Xiao, L., Sun, Y., Zhang, J., Ma, T., & He, L. (2021). A Survey of Human-in-the-loop for Machine Learning. <http://arxiv.org/abs/2108.00941>.
- [11] Van Der Stappen, A., & Funk, M. (2021). Towards Guidelines for Designing Human-in-the-Loop Machine Training Interfaces. *International Conference on Intelligent User Interfaces*, Proceedings IUI, 514–519. <https://doi.org/10.1145/3397481.3450668>.
- [12] McBride, N. (2021). Human in the loop. *Journal of Information Technology*, 36(1), 77–80. <https://doi.org/10.1177/0268396220946055>.
- [13] Steenwinckel, B., De Paepe, D., Vanden Haute, S., Heyvaert, P., Bentefrit, M., Moens, P., Dimou, A., Van Den Bossche, B., De Turck, F., Van Hoecke, S., & Ongenaes, F. (2021). FLAGS: A methodology for adaptive anomaly detection and root cause analysis on sensor data streams by fusing expert knowledge with machine learning. *Future Generation Computer Systems*, 116, 30–48. <https://doi.org/10.1016/j.future.2020.10.015>.
- [14] Guo, Y., Liao, W., Wang, Q., Yu, L., Ji, T., & Li, P. (2018). Multidimensional Time Series Anomaly Detection: A GRU-based Gaussian Mixture Variational Autoencoder Approach. *Proceedings of*

- Machine Learning Research, 95(2001), 97–112. <http://proceedings.mlr.press/v95/guo18a.html>.
- [15] Vilenski, E., Bak, P., & Rosenblatt, J. D. (2019). Multivariate anomaly detection for ensuring data quality of dendrometer sensor networks. *Computers and Electronics in Agriculture*, 162(November 2018), 412–421. <https://doi.org/10.1016/j.compag.2019.04.018>.
- [16] Cook, A. A., Misirli, G., & Fan, Z. (2020). Anomaly Detection for IoT Time-Series Data: A Survey. *IEEE Internet of Things Journal*, 7(7), 6481–6494. <https://doi.org/10.1109/JIOT.2019.2958185>.
- [17] Eredics, P., & Dobrowiecki, T. P. (2011). Data cleaning for an intelligent greenhouse. SACI 2011 - 6th IEEE International Symposium on Applied Computational Intelligence and Informatics, Proceedings, 293–297. <https://doi.org/10.1109/SACI.2011.5873017>.
- [18] Mehra, M., Saxena, S., Sankaranarayanan, S., Tom, R. J., & Veeramanikandan, M. (2018). IoT based hydroponics system using Deep Neural Networks. *Computers and Electronics in Agriculture*, 155(October), 473–486. <https://doi.org/10.1016/j.compag.2018.10.015>.
- [19] Castañeda-Miranda, A., & Castaño-Meneses, V. M. (2020). Internet of things for smart farming and frost intelligent control in greenhouses. *Computers and Electronics in Agriculture*, 176(May), 105614. <https://doi.org/10.1016/j.compag.2020.105614>.
- [20] Skelsey, P. (2021). Forecasting Risk of Crop Disease with Anomaly Detection Algorithms. *Phytopathology®*, PHYTO-05-20-018. <https://doi.org/10.1094/phyto-05-20-0185-r>.