International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 25 Issue 03, MAR, 2025

A CVAE-BASED ANOMALY DETECTION ALGORITHM FOR CYBER PHYSICAL ATTACKS FOR WATER DISTRIBUTION SYSTEMS

[1]Kaliki Seshaiah Babu, [2]A. Ravi Sankar MCA, M. Tech, M.Phil.,
[1]PG Scholar, Dept. of CSE, Srinivasa Institute of Technology and Science, Kadapa – 516001.
[2]Associate Professor & HOD, Dept. of CSE, Srinivasa Institute of Technology and Science, Kadapa – 516001.
[1]seshu645@gmail.com

ABSTRACT

Modern technologies adopt Internet of Things (IoT) devices to increase water management efficiency and enhance water quality services. However, the limitations of IoT devices, such as small sizes and poor security, weaken the Water Distribution System (WDS) security, and many attackers compromise the critical components of WDS. Cyber-physical attacks (CPAs) are considered one of the biggest challenges that decrease the security factors in WDS by disrupting normal operations and tampering with the critical data of the water system. For instance, an attacker can change the water pump's speed, disrupting the service. An attacker can also alter the data of water quality parameters to contaminate the water. It is important to propose solutions to increase security in the WDS and defend against CPAs and security threats. Although several intrusion detection methods were proposed in the literature to detect WDS CPAs, many issues still need solutions, such as detecting attacks with smaller false alarms, minimizing the time to disclose the attacks, determining the location of the compromising components, and recovering solutions for the attacked components. Therefore, this paper proposes a model based on a deep learning algorithm called a Conditional variational Autoencoder (CVAE) to disclose CPAs and mitigate their bad effects on WDS. The proposed method consists of a neural network, an encoder to compress data, and a decoder to decompress data. The objective goal is to minimize the reconstruction error between the encoded-decoded data and the initial data. We apply the CVAE on some well-known datasets. The experiment results show that our proposed CVAE method performs better than others. After analyzing the CVAE model with other existing models, we get the highest performance by reaching %98 accuracy.

I. INTRODUCTION

Water distribution systems (WDS) become smart by in-cluding modern technologies such as Internet of Things (IoT) devices [1], communication, remote control, intelligent decision-making, etc. These modern devices are parts of thebe-physical system, including smart water meters, pipes, valves, etc., that operate under both the physical aspect (such as a water system) and the cyber aspect (such as commune-cation networks, computing, and intelligence) [2]. Adopting new technologies improves the water system's efficiency and enhances the quality of services. However, the security levelof a WDS is decreased with increasing malicious events and cyber-physical attacks (CPAs) [3], [2]. CPAs expose the main components of the WDS and the communication links between these components. For example, a CPA can compromise a pump to prevent providing water. Also, a Cepacian change the data of components to disrupt the operation, such as changing the water level in a water tank to overload waste the water [3]. Additionally, some CPAs target the water quality sensors to contaminate or poison the water [4], which threatens customers' safety threats, vulnerabilities, and attacks.

1.1 KNN Classifier Algorithm

K-nearest neighbor method can be used for both regression and classification predictive problems. This method helps in interpret output, calculate time and predictive power. The Machine learning techniques are used in various fields. KNN is also one of the machine learning methods. This is also called as method of sample-based learning. This will contain the data of past datasets and can be used while predicting the new datasets. This will apply function called as distance function like Manhattan or Euclidean distance. This can be used to compute distance from samples to all other training samples. It calculates the target value for new samples. The target vale will be the weighted sum of target values of the k nearest neighbours. The valve of K can be directly proportional to the prediction. Whenever the valve of K is small this indicates there is high variance and there is low bias. If the valve of the K is larger than this indicates that there is low variance and high bias. The main advantage of this KNN is it does not require any training or the optimization. This KNN uses data samples when predicting the new datasets. Hence it is having higher complexity and also more time consumption.

This work represents a review of K-NN technique for the early prediction of food recommendation. K-NN analysis is used for predicting the unknown parameter from the known parameters. In this work we are considering vitamins as input parameters which are the main parameters to be considered for a good food recommendation, although there are many other factors that can be considered [17]. The unknown value of vitamis can be predicted from the nearest known values of the nearest neighbors by calculation of Euclidean distance between them. Then we would be able to predict type of food for given vitamin parameters. To measure the distance between points in a feature space, various distance functions can be used, in which the Euclidean distance function is the most widely used one[18]. Let p and q are represented as feature vectors. To calculate the distance between p and q, the Euclidean metric is generally used by if a=(a1, a2) and b=(b1,b2) then the distance is given

1.2 Classifications Algorithms

Onto the part you've probably been waiting for all this time: training machine learning algorithms. To be able to test the performance of our algorithms, I first performed an 80/20 train-test split, splitting our balanced data set into two pieces. To avoid overfitting, I used the very common resampling technique of k-fold cross-validation. This simply means that you separate your training data into k parts (folds) and then fit your model on k-1 folds before making predictions for the kth hold-out fold. You then repeat this process for every single fold and average the resulting predictions.

1.3 Logistic regression

This is a classification function that uses class for building and uses a single multinomial logistic regression model with a single estimator. Logistic regression usually states where the boundary between the classes exists, also states the class probabilities depend on distance from the boundary, in a specific approach. This moves towards the extremes (0 and 1) more rapidly when data set is larger. These statements about probabilities which make logistic regression more than just a classifier. It makes stronger, more detailed predictions, and can be fit in a different way; but those strong predictions could be wrong. Logistic regression is an approach to prediction, like Ordinary Least Squares (OLS) regression. However, with logistic regression, prediction results in a dichotomous outcome [13]. Logistic regression is one of the most commonly used tools for applied statistics and discrete data analysis. Logistic regression is linear interpolation.

II. SYSTEM ANALYSIS

The Systems Development Life Cycle (SDLC), or Software Development Life Cycle in systems engineering, information systems and software engineering, is the process of creating or altering systems, and the models and methodologies that people use to develop these systems. In software engineering the SDLC concept underpins many kinds of software development methodologies.

2.1 Existing System

Many attack detection methods are proposed in the literature[2]. The authors in [28] propose a MATLAB toolbox calledepanet CPA to design CPA attacks and simulate the negative impacts of these attacks on the WDS. However, the tool does not provide any detection method to detect CPAs. The authors in [29] propose a method to reduce the vulnerability of CPAs by applying a logic graph algorithm.

The idea of the logic graph algorithm is to model WDS components asa graph, including interconnection between the sensors data and actuators data to compute the control function values for all subgraphs and compare it with existing control function values to discover the attacks. However, this work has the drawback of not considering IoT security, and it can only be used for a simple water network. The detection method in[30] adopts the Autoencoder algorithm (AE) to detect CPAs.The AE trains data and detects anomalies based on a thresh-old. The result shows that the AE could detect attacks but could not recognize the locations of the attacked components. The authors in [31] combine the following three algorithms to detect CPAs.

The first algorithm adopts a statical algorithm to detect outliers; the second algorithm trains data to define the attacks by applying an Artificial Neural Network (ANN), and the third algorithm adopts the principal components analysis(PCA) to classify the data to normal and outliers.

a) Disadvantages of Existing System

- Limited ability to capture intricate relationships in vast, high-dimensional datasets.
- Susceptible to high false positive rates, resulting in user inconvenience and additional verification costs.
- Difficulty adapting to new fraud tactics, requiring frequent retraining with new data.
- High dependency on feature engineering, making it labor-intensive and prone to missing subtle patterns.
- Less effective in real-time scenarios due to limited scalability and adaptability.

2.2 Proposed System

Our proposed method adopts the Conditional Variational Autoencoder (CVAE) based on the Variational Autoencoder (VAE) The autoencoder is one type of neural network algorithm and can reduce high-dimension data to lower-dimension data by compressing data. It is an unsupervised method that learns from datasets efficiently without labeling data [42]. It includes three parts, as shown in Figure: an encoder, latent space, and a decoder [43]. The first part is called an encoder, which compresses the input data by converting input data from high dimensions to the lowest dimensions. The second part is a latent space, representing the space of the reduced dimensions of input or compressed input data. Also, the latent space can be called a bottleneck architecture, which contains the maximum numbers of the compression input data [44]. The third part is called a decoder, which reconstructs the output data from the latent space and back to the approximate input data.

b) Advantages of Proposed System

- Captures complex relationships in large, high-dimensional financial data using neural networks.
- Reduces false positives through advanced pattern recognition, enhancing detection accuracy.
- The distributed framework enables real-time fraud detection and faster data processing.
- Highly adaptable to evolving fraud tactics without extensive retraining.
- Decreases dependency on manual feature engineering, resulting in more effective and autonomous detection.

III. SYSTEM DESIGN

3.1 System architecture

A system architecture or systems architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system.

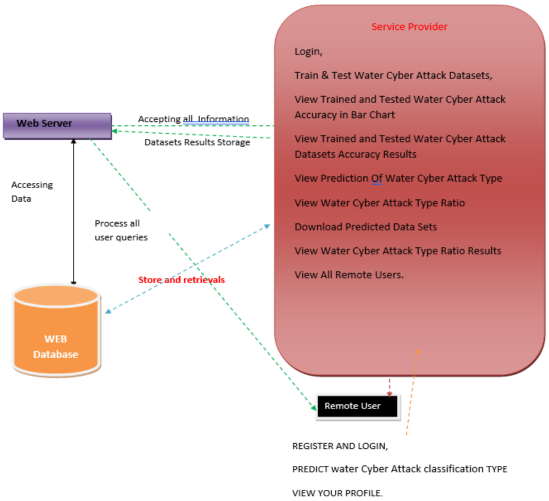

Service Provider
Login,
Train & Test Water Cyber Attack Datasets,
View Trained and Tested Water Cyber Attack Accuracy in Bar Chart
View Trained and Tested Water Cyber Attack Datasets Accuracy Results
View Prediction Of Water Cyber Attack Type
View Water Cyber Attack Type Ratio
Download Predicted Data Sets
View Water Cyber Attack Type Ratio Results
View All Remote Users.
Web Server
Accepting all Information
Datasets Results Storage
Accessing Data
Process all user queries
Store and retrievals
WEB Database
Remote User
REGISTER AND LOGIN,
PREDICT water Cyber Attack classification TYPE
VIEW YOUR PROFILE.

Figure – 1: System Architecture

a) 3-Tier Architecture

The three-tier software architecture (a three-layer architecture) emerged in the 1990s to overcome the limitations of the two-tier architecture. The third tier (middle tier server) is between the user interface (client) and the data management (server) components. This middle tier provides process management where business logic and rules are executed and can accommodate hundreds of users (as compared to only 100 users with the two-tier architecture) by providing functions such as queuing, application execution, and database staging.

The three-tier architecture is used when an effective distributed client/server design is needed that provides (when compared to the two tier) increased performance, flexibility,maintainability, reusability, and scalability, while hiding the complexity of distributed processing from the user. These characteristics have made three-layer architectures a popular choice for Internet applications and net-centric information systems.

IV. IMPLEMENTATION

a) Design

The software system design is produced from the results of the requirements phase. Architects have the ball in their court during this phase and this is the phase in which their focus lies. This is where the details on how the system will work is produced. Architecture, including hardware and software, communication, software design (UML is produced here) are all part of the deliverables of a design phase.

b) Implementation

Code is produced from the deliverables of the design phase during implementation, and this is the longest phase of the software development life cycle. For a developer, this is the main focus of the life cycle because this is where the code is produced. Implementation my overlap with both the design and testing phases. Many tools exists (CASE tools) to actually automate the production of code using information gathered and produced during the design phase.

V. TESTING

Testing is the process where the test data is prepared and is used for testing the modules individually and later the validation given for the fields. Then the system testing takes place which makes sure that all components of the system property functions as a unit. The test data should be chosen such that it passed through all possible condition. The following is the description of the testing strategies, which were carried out during the testing period.

During testing, the implementation is tested against the requirements to make sure that the product is actually solving the needs addressed and gathered during the requirements phase. Unit tests and system/acceptance tests are done during this phase. Unit tests act on a specific component of the system, while system tests act on the system as a whole.

So in a nutshell, that is a very basic overview of the general software development life cycle model. Now let’s delve into some of the traditional and widely used variations.

5.1 System Testing

Testing has become an integral part of any system or project especially in the field of information technology. The importance of testing is a method of justifying, if one is ready to move further, be it to be check if one is capable to with stand the rigors of a particular situation cannot be underplayed and that is why testing before development is so critical. When the software is developed before it is given to user to user the software must be tested whether it is solving the purpose for which it is developed. This testing involves various types through which one can ensure the software is reliable. The program was tested logically and pattern of execution of the program for a set of data are repeated. Thus the code was exhaustively checked for all possible correct data and the outcomes were also checked.

5.2 Module Testing

To locate errors, each module is tested individually. This enables us to detect error and correct it without affecting any other modules. Whenever the program is not satisfying the required function, it must be corrected to get the required result. Thus all the modules are individually tested from bottom up starting with the smallest and lowest modules and proceeding to the next level. Each module in the system is tested separately. For example the job classification module is tested separately. This module is tested with different job and its approximate execution time and the result of the test is compared with the results that are prepared manually. Each module in the system is tested separately. In this system the resource classification and job scheduling modules are tested separately and their corresponding results are obtained which reduces the process waiting time.

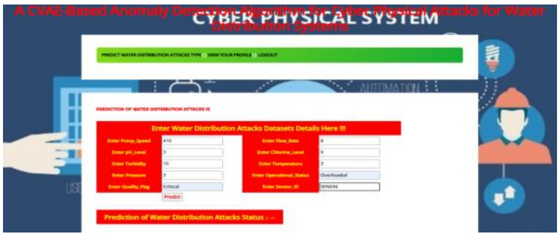
5.3 Integration Testing

After the module testing, the integration testing is applied. When linking the modules there may be chance for errors to occur, these errors are corrected by using this testing. In this system all modules are connected and tested. The testing results are very correct. Thus the mapping of jobs with resources is done correctly by the system.

5.4 Acceptance Testing

When that user fined no major problems with its accuracy, the system passes through a final acceptance test. This test confirms that the system needs the original goals, objectives and requirements established during analysis without actual execution which elimination wastage of time and money acceptance tests on the shoulders of users and management, it is finally acceptable and ready for the operation.

VI. OUTPUT SCREENS


A CVAE-Based Anomaly Detection Algorithm for Cyber Physical Attacks for Water Distribution Systems
CYBER PHYSICAL SYSTEM
PREDICT WATER DISTRIBUTION ATTACKS TYPE || VIEW YOUR PROFILE || LOGOUT
AUTOMATION
Predict

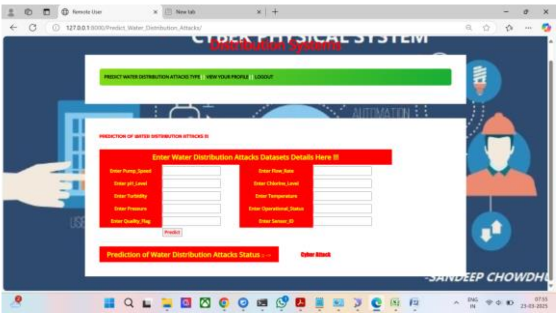
PREDICTION OF WATER DISTRIBUTION ATTACKS !!!

Enter Water Distribution Attacks Datasets Details Here !!!

Enter Pump_Speed	410	Enter Flow_Rate	8
Enter pH_Level	3	Enter Chlorine_Level	9
Enter Turbidity	10	Enter Temperature	3
Enter Pressure	3	Enter Operational_Status	Overloaded
Enter Quality_Flag	Critical	Enter Sensor_ID	SEN036

Prediction of Water Distribution Attacks Status :-->

Figure -2: Home Page

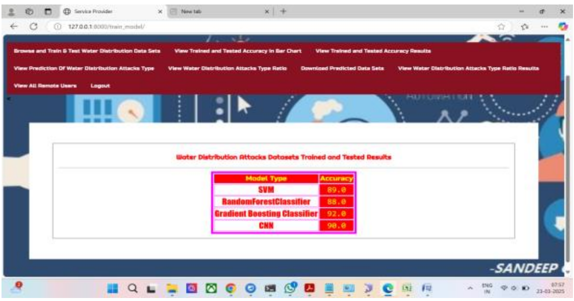


PREDICTION OF WATER DISTRIBUTION ATTACKS !!!

Enter Water Distribution Attacks Datasets Details Here !!!

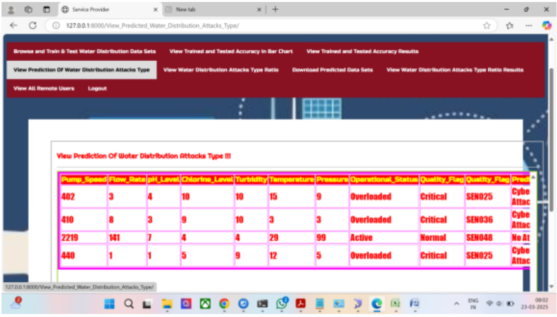
Prediction of Water Distribution Attacks Status :--> Cyber Attack

Figure -3: Distribution System


Service Provider
New tab
127.0.0.1:8000/train_model/
Browse and Train & Test Water Distribution Data Sets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Prediction Of Water Distribution Attacks Type
View Water Distribution Attacks Type Ratio
Download Predicted Data Sets
View Water Distribution Attacks Type Ratio Results
View All Remote Users
Logout
-SANDEEP
07:57
23-03-2025

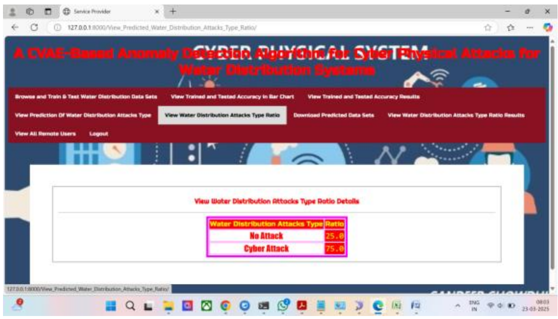
Water Distribution Attacks Datasets Trained and Tested Results

Model Type	Accuracy
SVM	89.0
RandomForestClassifier	88.0
Gradient Boosting Classifier	92.0
CNN	90.0


Service Provider
New tab
127.0.0.1:8000/View_Predicted_Water_Distribution_Attacks_Type/
Browse and Train & Test Water Distribution Data Sets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Prediction Of Water Distribution Attacks Type
View Water Distribution Attacks Type Ratio
Download Predicted Data Sets
View Water Distribution Attacks Type Ratio Results
View All Remote Users
Logout
08:02
23-03-2025

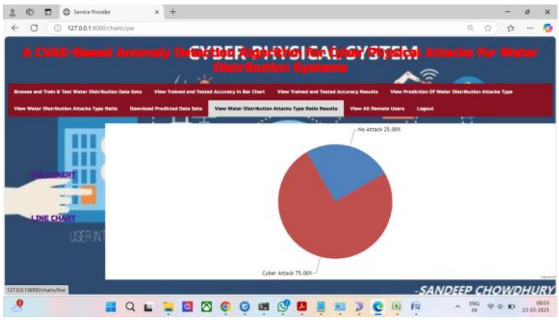
View Prediction Of Water Distribution Attacks Type !!!

Pump_Speed	Flow_Rate	pH_Level	Chlorine_Level	Turbidity	Temperature	Pressure	Operational_Status	Quality_Flag	Quality_Flag	Predi
402	3	4	10	10	15	9	Overloaded	Critical	SEN025	Cybe Attac
410	8	3	9	10	3	3	Overloaded	Critical	SEN036	Cybe Attac
2219	141	7	4	4	29	99	Active	Normal	SEN048	No At
440	1	1	5	9	12	5	Overloaded	Critical	SEN025	Cybe Attac


Service Provider
127.0.0.1:8000/View_Predicted_Water_Distribution_Attacks_Type_Ratio/
A CVAE-Based Anomaly Detection Algorithm for Cyber Physical Attacks for Water Distribution Systems
CYBER PHYSICAL SYSTEM
Browse and Train & Test Water Distribution Data Sets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Prediction Of Water Distribution Attacks Type
View Water Distribution Attacks Type Ratio
Download Predicted Data Sets
View Water Distribution Attacks Type Ratio Results
View All Remote Users
Logout
08:03
23-03-2025

View Water Distribution Attacks Type Ratio Details

Water Distribution Attacks Type	Ratio
No Attack	25.0
Cyber Attack	75.0


Service Provider
127.0.0.1:8000/charts/pie
A CVAE-Based Anomaly Detection Algorithm for Cyber Physical Attacks for Water Distribution Systems
CYBER PHYSICAL SYSTEM
Browse and Train & Test Water Distribution Data Sets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Prediction Of Water Distribution Attacks Type
View Water Distribution Attacks Type Ratio
Download Predicted Data Sets
View Water Distribution Attacks Type Ratio Results
View All Remote Users
Logout
PIE CHART
LINE CHART
No Attack 25.00%
Cyber Attack 75.00%
-SANDEEP CHOWDHURY
08:03
23-03-2025

VII. CONCLUSION

IoT and smart devices increase the ability of WDSs to serve customers efficiently. However, the limitations of advanced technology, such as small memory and low capability, reduce the security and privacy of WDSs. These limitations attract attackers to disrupt normal operations and alter the critical data. To address the security issues and mitigate CPAs, we propose a model based on the CVAE algorithm to detector mitigate CPAs. CVAE is a deep learning algorithm that extends the VAE algorithm to control data. CVAE can model complex structures based on probabilistic inference. The proposed model shows high efficiency in defending against with high performance. The proposed model based on the CVAE algorithm should be further extended to enhance and increase the security of the WDSs.

7.1 Future Enhancements

Future advancements could include the integration of deep learning techniques for even more accurate anomaly detection, the use of advanced threat intelligence for predictive analysis, and the development of autonomous response systems. Additionally, extending the system to hybrid cloud environments and enhancing user privacy in behavioral monitoring could further improve security resilience.

REFERENCES

[1] S. Diaz, A. Molano, C. Erazo, and J. C. Monroy, “Wqms: water qualitymonitoring station for iot,” Int. J. Sens. Netw., vol. 35, no. 2, pp. 79–87,2021.

[2] A. A. Abokifa, A.M.ASCE, K. Haddad, C. Lo, and P. Biswas, “Real-timeidentification of cyber-physical attacks on water distribution systems viamachine learning–based anomaly detection techniques,” Journal of WaterResources Planning and Management, vol. 145, no. 1, 2018.

[3] C. M. Ahmed, V. R. Palleti, and A. P. Mathur, “Wadi: A water distributiontestbed for research in the design of secure cyber physical systems,”in Proceedings of the 3rd International Workshop on Cyber-PhysicalSystems for Smart Water Networks, ser. CySWATER ’17. New York, NY,USA: Association for Computing Machinery, 2017, p. 25–28. [Online].Available: https://doi.org/10.1145/3055366.3055375

[4] S. Adepu and A. Mathur, “An investigation into the response of a watertreatment system to cyber attacks,” in 2016 IEEE 17th InternationalSymposium on High Assurance Systems Engineering (HASE), 2016, pp.141–148.

[5] R. Khelif, M. Kharrat, and M. Abid, “A novel design of the multi-wire-based solution for water leak detection and localisation in buried pipes,”Int. J. Sens. Netw., vol. 39, no. 2, pp. 83–92, 2022

[6] Cho, Byeong-joo, Jang-ho Yun, and Kyeong-ho Lee. "Study of effectiveness for the network separation policy of financial companies." Journal of the Korea Institute of Information Security & Cryptology 25.1 (2015): 181-195.

[7] Bagui, Sikha, et al. "Detecting reconnaissance and discovery tactics from the MITRE ATT&CK framework in Zeek Conn Logs using Spark’s machine learning in the big data framework." Sensors 22.20 (2022): 7999.

[8] Anjum, Md Monowar, Shahrear Iqbal, and Benoit Hamelin. "Analyzing the usefulness of the DARPA OpTC Dataset in cyber threat detection research." Proceedings of the 26th ACM Symposium on Access Control Models and Technologies. 2021.

[9] API Reference - scikit-learn 1.3.2 documentation [Online]. Available: https://scikit-learn.org/stable/modules/classes.html, Accessed on: Dec. 1, 2023.

[10] Log Files - Book of Zeek [Online]. Available: https://docs.zeek.org/en/master/script-reference/log-files.html, Accessed on: Dec. 1, 2023.