**SUPPLY CHAIN ANOMALY DETECTION**

*Project report Submitted*

*to*

**MANIPAL ACADEMY OF HIGHER EDUCATION**

*For Partial Fulfillment of the Requirement for the*

*Award of the Degree*

*Of*

**Bachelor of Technology**

*in*

**Computer and Communication Engineering**

*by*

**NALLAMILLI NAGA VENKATA REDDY-210953334**

**VARIKUTI MADHURIMA-210953314**

**VURA SAI NIKHIL-210953204**

*Under the guidance of*

Dr. Chethan Sharma Dr. Raghavendra s

Associate professor-senior scale Associate professor-senior scale

Department of I&CT Department of I&CT

Manipal Institute of Technology Manipal Institute of Technology

****Manipal, Karnataka, India Manipal, Karnataka, India

**November 2023**

**Table of contents:**

**Chapter 1. Introduction……………………………… 3**

**Chapter 2. Literature survey………………………… 4**

**Chapter 3. Methodology……………………………… 8**

**Chapter 4. Results and discussions….……………. …17**

**Chapter 5. Conclusion………………………………... 29**

**References………………………………… 30**

**Chapter-1**

**Introduction:**

In today’s context, the supply chain plays a significant role in moving goods and services from manufacturers to end users. However, this supply chain is vulnerable to disruptions and there are unpredictable changes in demand. Identifying anomalies or unusual patterns in the supply chain is important for companies to maintain operational quality, reduce risk, and increase overall resilience Anomaly detection is the detection of data points or patterns that deviate significantly from the expected or normal behavior of a system or data set. In supply chain management, anomaly identification can be a powerful tool to anticipate and mitigate problems before they turn into serious issues.

**Current Practices**:

Looking at how things are done now in managing supply chains, we see a trend towards using more data to make decisions. The techniques we're using stand out for being good at finding anomalies in different situations. This project aims to use these techniques to solve the challenge of finding odd things in complex supply chain data.

**Applications Related to the Project:**

Finding anomalies is like discovering hidden problems. In the data about how things are priced for shipping, we want to make sure everything is fair and clear, with no overcharging or underpricing. In the logistics data, we're looking for signs of delays, disruptions in transportation, or strange patterns in how things are stored. If we find these problems early, we can fix them and make the whole system work better.

**Motivation:**

The motivation for the project "Supply Chain Anomaly Identification Using Outlier Analysis" is based on the critical need for businesses to adapt to the challenges posed by the dynamic and interconnected nature of today's supply chain thus helping to build a simpler, more cost-effective, adaptive supply chain design.

**Chapter-2**

**Literature survey**

Baumgartner et al. [1]

This paper gives an approach to simulate and analyze the different anomalies in Boxed Beef supply chain routes by using sensors, simulating different scenarios of supply chain anomalies, and distinguishing between different types of anomalies. Simulation is used to recreate the different scenarios to detect the anomalies by using temperature sensors, trip durations and cooling requirements. It uses logic-based language which can potentially observe the anomalies and determine the type of anomaly. The scenarios used in this paper are fit to be tested on to get accurate data and results from the supply chain. The Algorithm can be applied to other similar types of products to remove anomalies in their supply chain by making some changes to the algorithm and the scenarios by studying the model of the supply system. The potential advancements in the algorithms to make the algorithm useful for wide range of applications by increasing its functionalities. The algorithm increases the quality of Food and reducing the wastage from the problems faced in the supply chain.

Jiménez-Carmelo et al. [2]

This paper discusses about how to tackle food frauds using unique identification labels for the food as it helps in food traceability and recording of data about products and their respective batches. Unique ID and analysis using artificial intelligence of the data from the Unique ID reveals interesting details and patterns about the product to make sure the originality if the product is not tampered and Frauds are reduced. intrinsic approach in which markers on the products which change throughout the supply chain. extrinsic approach, the use of QR codes, barcodes which are digitally scanned and retrieve data ensuring data security and privacy and reducing the Counterfeit products. The use case of this algorithm in different scenarios in supply chain can give different results, some can give excellent results, and some cannot. This paper focuses attention on analysis on spatial and temporal data which is generated in traceability system can be an early sign to identify and mark the anomalies that are suspected in supply chain.

Sahin et al. [3]

This paper aims to explain about an algorithm which is an iterative optimized algorithm. It helps in transporting the container in a distribution system. Dynamic nature of maritime sector due to changing in weather, financial conditions and risks leads to uncertainties. This paper suggests a model that can be a perfect solution to the dynamic change in maritime sector. Shortest path algorithm like Dijkstra’s algorithm is used in the optimized techniques to decrease the cost, risk and to achieve high level performance.Calculating weights for optimizing risks, cost, and performance in the networks. A fuzzy analytical hierarchy procedure is used as a multicriteria decision method. To achieve this result. But due to its modular nature it can adapt to new technologies, can be improved by embedding new algorithms to work faster. Maritime sector needs an optimized model which is dynamic and can withstand unforeseen challenges. The algorithm can be applied to a more complex dynamic fuzzy environments like incorporating real-time big data and adapting to changing conditions in maritime supply chain.

Yeboah-Ofori et al. [4]

This paper reveals an innovative approach to detecting supply chain attacks in0 CPS domain. BBN It is a graphical model which uses Bayesian probability theory to understand dependencies among a set of variables. Collect data related to the supply chain operations and then clean and preprocess the data to remove inconsistencies. Determine the relationships between variables and design by specifying nodes for each variable and edges, Conditional probability tables. Probabilistic Modelling: Provides a comprehensive view of the attack scenario. Performance can vary depending on many factors like Complexity of the model, detection speed, data quality, and resource utilization. The performance will depend on how well these challenges are managed. BBN can enhance CPS security and it is the best approach to detect cyber-attacks. When developed wisely, BBNs serve as a valuable tool.

Chukkapalli et al. [5]

In this paper, we discuss the framework which is used to preserve smart farm owner’s privacy. The leakage of data from the corporation will have a major impact on the farms which cause a loss in economics. Data perturbation with white Gaussian noise: It is a privacy-protecting technique, and this technique is widely used for data transformation while processing the signal for adding random noise to the original data or electronic devices. This framework consists of 4 modules which include data collection and exchange ontology, knowledge ledger The outcomes of the experiments demonstrated that it can be analysed through an anomaly detection model It can also be measured through data split and models used. Its performance on transformed data is slightly worse than that of non-transformed data. It is proved to be a promising approach and beneficial to and integrity of individual smart farm data.

Dimitrios Tychalas et al. [6]

The paper focuses on the growing threat of data theft in Industrial Control Systems (ICS) via the supply chain, using a firmware modification that unites the air gap and data is transferred from memory to analogue devices. By short-circuiting specific components in the aimed target device by making use of the Device Tree and lies exclusively in the devices connected externally, totally hidden from the main CPU. Implements the attack on an industrial Programmable Logic Controller, The paper fails to furnish a comprehensive evaluation of the attack's impact on the security and integrity. In terms of CPU performance and speed at which DMA is being transferred the introduced attack vector does not demonstrate any perceivable overhead. It emphasizes the crucial requirement for specialized strategies adopted to industrial control systems to quickly spot and rectify these menaces, while also stressing the significance of robust supply chain security measures.

Zunic et al. [7]

A progressive multi-stage anomaly detection set of rules with actual-global software within the context of supply chain control. In this process, it is very important to verify the accuracy of data which are used as data errors can result in significant additional costs it mainly applies to companies with similar services and processes. this paper gives individual processes that can be adapted for warehouse applications. This algorithm is implemented in some of the big and top distributed companies, showing its practical applicability and effectiveness, and this algorithm operates on the smart warehouse management system. Median-based improves the accuracy of anomaly detection. This algorithm combines online and offline parts and uses clusters and median. The importance of detection in information systems, its potential for cost savings, and an increase in the belief in data accuracy.

Vu et al. [8]

The main agenda of this article is to use the repositories, such as GitHub. This proceeds towards identifying threats and attacks through lightweight analysis. This will decrease the review effort by 97 percent and the processing time is speed. This procedure involves studying software artifacts from the repositories and analysing dispersed artifacts in different package repositories. This procedure includes various artifacts from different Python packages. Upcoming work includes making a classifier among benign and malicious administration. Processing speed will be extremely fast and speedy, which is one of the median times for separate artifacts. The algorithm will have speed processing techniques with a median time for processing and the upcoming work involves work that consists of making a classifier among benign and malicious administration. The lightweight analysis process will be because of distributed artifacts and repositories. This procedure will give fast processing in less than 33 sec. and finally, it is an efficient approach for on-the-fly detection of the software supply and finding suspicious activity.

Rejeb et al. [9]

This article focuses on the merging of IoT and blockchain technology. There are a few existing topics like traceability, efficiency, effectiveness and auditing, immutability, security, and scalability. The article has 6 research topics that trace the key features that are mentioned in the problem statement. Blockchain makes possible safe communication and honesty of exchange transactions, which will allow direct transactions with the machines through good contracts, it allows to call back the unsecured goods, and events. Product attribute. It refines the visibility of the product. And betters the traceability. live monitoring is one of the main advantages of this. technologies can improve final traceability. IoT solutions can improve operational costs. Both technologies have modern supply chains, and it will increase efficiency and effectiveness. Combination of IoT and blockchain will give better solutions to the problems. In both technologies, blockchain can face confrontations in supply chain management, and IoT enhances operational performance which includes traceability. It has a great capability for supply chain management.

Tsoukas et al [10]

The main theme of this article is to define the need for new approaches to control quality parameters. The blockchain and the TinyML the initial thing is to maintain data integrity it secures data tampering and gives clarity for all actors in the supply chain, and the next thing is to ease malicious behaviour. Using blockchain technology to pledge the purity of collected information. A technique called the self-sovereign identity approach is used for all actors to minimize single points of failures. The transparency levels are extremely elevated than related to the remaining and the quality parameters are very remarkable. anomaly detection assists in recognizing irregularities in the supply chain, and data integrity. It reviews performance through False Negative Rate [FNR] and False Positive Rate [FPR], and precision, and so forth.

The paper showcases a novel approach to enhancing food supply chain security by integrating blockchain technology and TinyML. The system encourages disclosure in the food industry. TinyML anomaly detection makes it easier to report misconduct to actors in the supply chain.

**Chapter-3**

**Methodology:**

The data we have chosen is supply chain shipment pricing data which contains twenty-four attributes with a huge amount of 10,325 tuples which can be used to analyze the dataset and find out the anomalies and outliers and make those predictions. The attributes are as follows:

* Id: it is a unique attribute which is different from everyone it is used for tracking and for referring purposes.
* Project code: it is a code for projects which is assigned for identification purpose.it is used for categorizing and organizing projects.
* Pq#: pq# it is prequalification number which is a unique number associated with prequalification process.
* Po/so#: po stands for purchase order, so stands for sales order. It is used for transactions between buyer and seller.
* Asn/dn#: asn stands for advance shipping notice and dn stands for delivery note number. These are used in the shipping process to notify about incoming shipments.
* Country: it is associated with the project, it indicates location.
* Managed by the person who is responsible for managing the project.
* Fulfill via method which is responsible for fulfilling project needs.
* Vendor inco term: it represents responsibilities of buyer and seller in international trade transactions.
* Shipment mode: mode of transportation used for shipping the products.
* pq first sent to client date: date when prequalification information was first sent to client.
* po sent to vendor date: the date when the purchased order was sent to vendor.
* scheduled delivery date: planned date for delivery of products.
* Delivered to client date: the actual date when product need to be delivered.
* Delivery recorded date: the delivery information when it was recorded.
* Product group: the category group which product belongs.
* Sub classification: subcategory of a product
* Vendor: the supplier who is providing the products
* Item description: a detailed description of project
* Molecular/test type: the specific molecule or type associated with project.
* Brand: brand name of product
* Dosage: amount of dose of product.
* Dosage form: form in which dosage is administered.
* Unit of measure (per pack): the unit used to measure quantity of product.
* Line-item quantity: quantity of the product ordered.
* line-item value: total values which are associated with line item.
* pack price: price per pack of product
* unit price: price per unit of product
* manufacturing site: location where product is manufactured.
* first line designation: designates the status of first line project.
* weight(kilograms): weight of products in kilograms.
* Freight cost (USD): cost associated with shipping or freight in u.s dollars.
* Line-item insurance (USD): insurance cost associated with line-item in us dollars.

**Non continuous attributes:** [“id”, ”project code”, “pq#”, “ po/so #”, “asn/dn #”, “country”, “managed by”, “fulfill via”, “vendor inco term”, “shipment mode”, “pq first sent to client date”, “po sent to vendor date”, “scheduled delivery date”, “delivered to client date”, “delivery recorded date”, “product group”, “sub classification”, “vendor”, “item description”, “molecule/test type”, “brand”, “dosage”, “dosage form”, “unit of measure (per pack)”, “manufacturing site”, “first line designation”, “freight cost (usd)”]

**Continuous attributes**: [“line\_item\_quantity,” “unit\_price”, “pack\_price” , “line\_item\_value”, “line\_item\_insurance”, “sub\_classification”, “dosage”, “weight”]

**Data cleaning:**

We can find the null values in the dataset using **df. isnull().sum()** which is used to find no of null values in our data.

For the null values we found we are replacing them with a value by finding mean of a value **mean\_value=df["line\_item\_insurance"].mean()mean\_value.**

This function replaces all the null values with mean of line\_item\_insurance values:

**df["line\_item\_insurance"].fillna(mean\_value,inplace=True)**

again, it checks if there are any null values:

**df.isnull().sum()**

a new data frame ‘df1’ is created by removing null values:

**df1=df.dropna()**

now it checks for null values in data frame:

**df1.isnull().sum()**

so, using **df1.info()** we can confirm that any null values are there are not.

so, we sent that data into new data file which has cleaned values and cleaned data.

**1.Local outlier factor (LOF):**

Lof is a density-based algorithm which identifies anomalies by comparing the local density of data points to their neighbours.

Lof is a unsupervised learning algorithm

It calculates the score for each data point based on ratio of local density of that point to local density of its neighbours.

Anomalies can be identified as the points with lower density compared to neighbours.

Local Outlier Factor(lof) focuses on identifying unusual or anomalies data points.

Step-1: import libraries

Step-2:data loading:

We have loaded dataset from a file named “cleaned\_data\_1.csv”

Step-3: selection of features:

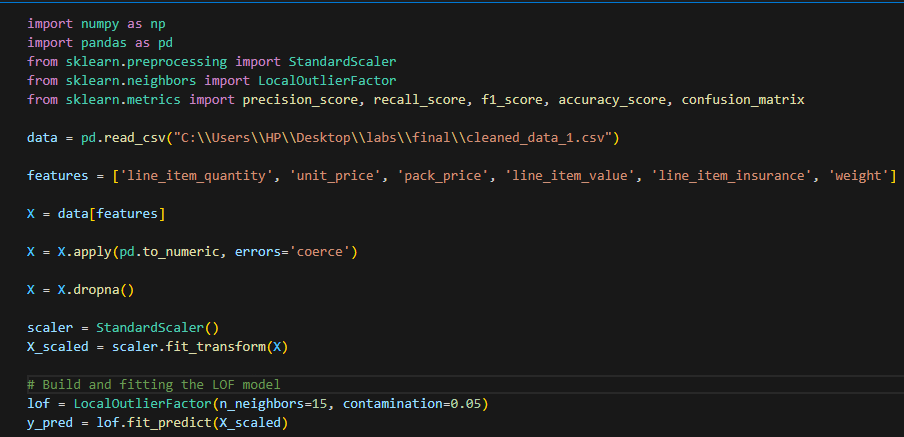
: [“line\_item\_quantity,” “unit price”, “pack price”, “line\_item\_value”, “line\_item\_insurance”, “weight”]

step 4: scaling:for facilitating distance-based calculations for the selected features we are using “StandardScaler” function from scikit-learnthe scaled information has been fitted into Scaled. step 5: LOF algorithm**model configuration**:

lof algorithm was instantiated with following parameters they are n\_neighbours, contamination.

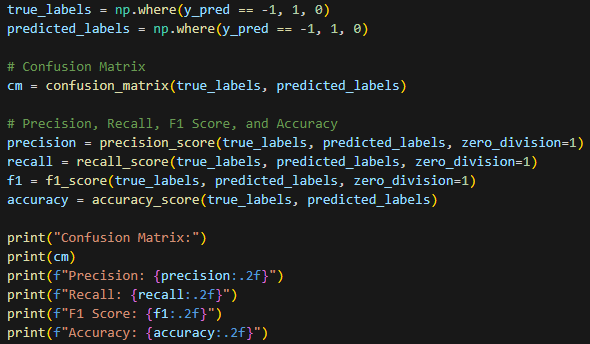
model fitting:

the scaled data was used to fit the LOF model enabling the algorithm to learn that patterns within feature space.



step 6: model evaluation:

lof predictions are used as true labels (-1 for anomalies, 1 for normal) and zero division is used to manage potential divisions by zero in precision, recall, f1 score.



**2.ISOLATION FOREST:**

Isolation forest is an ensemble-based algorithm that isolates anomalies by constructiong a random forest of isolation trees. Isolation forest is an unsupervised learning algorithm. It will measure no. of points needed to isolate a data point as a measure of its anomalies. Anomalies are isolated more quickly than normal points.

Step-1: import required libraries.

Step-2:data loading

We have loaded dataset from a file named “cleaned\_data\_1.csv”

Step-3: selection of features:

: [“line\_item\_quantity,” “unit\_price”, “pack\_price, “line\_item\_value”, “line\_item\_insurance”, “weight”]

Step-4:data cleaning

Convert the selected columns to numeric format, managing any conversion errorsand dropping rows with missing values.

Step-4: feature scaling

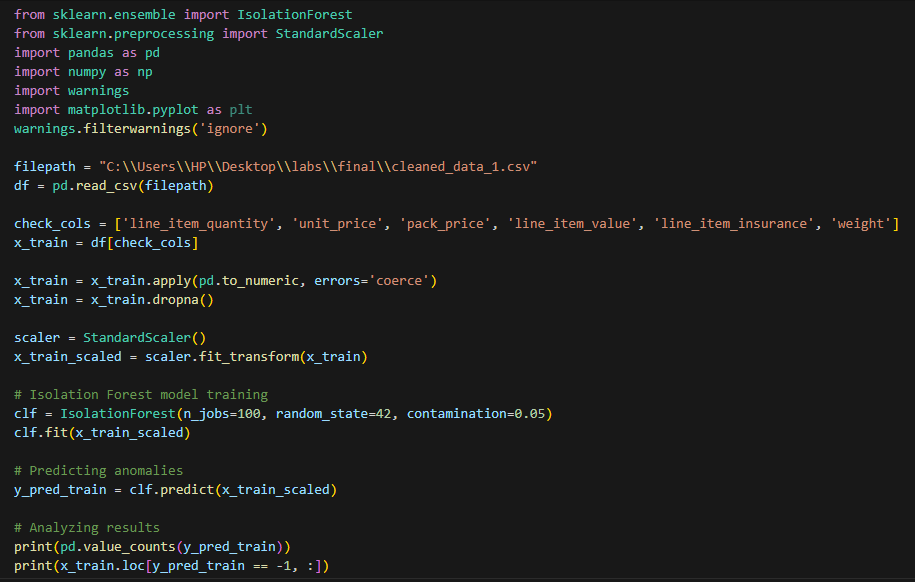
Applied feature scaling using “standardscaler” to ensure uniformity in the given features.

Step-5: isolation forest model training:

Train the isolation forest model based on scaled data and set the model parameters that are n\_jobs, random state, contamination rate, random state and fit the x\_train\_scaled.

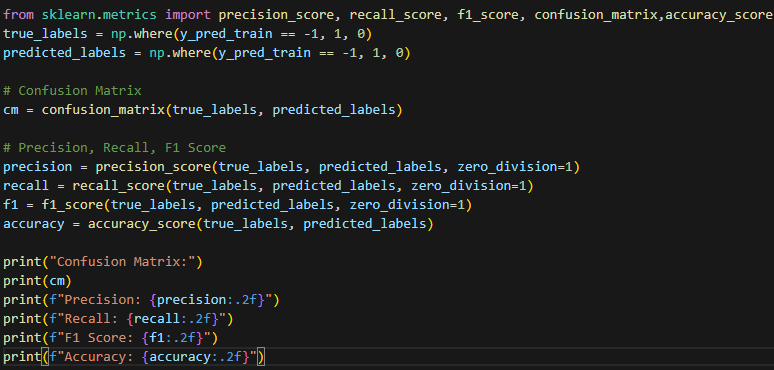
Step-7: predict anomalies:

Predicting anomalies in training data using isolation forest model



Step-8: model evalution:

lof predictions are used as true labels (-1 for anomalies, 1 for normal) and zero division is used to oversee potential divisions by zero in precision, recall, f1 score.



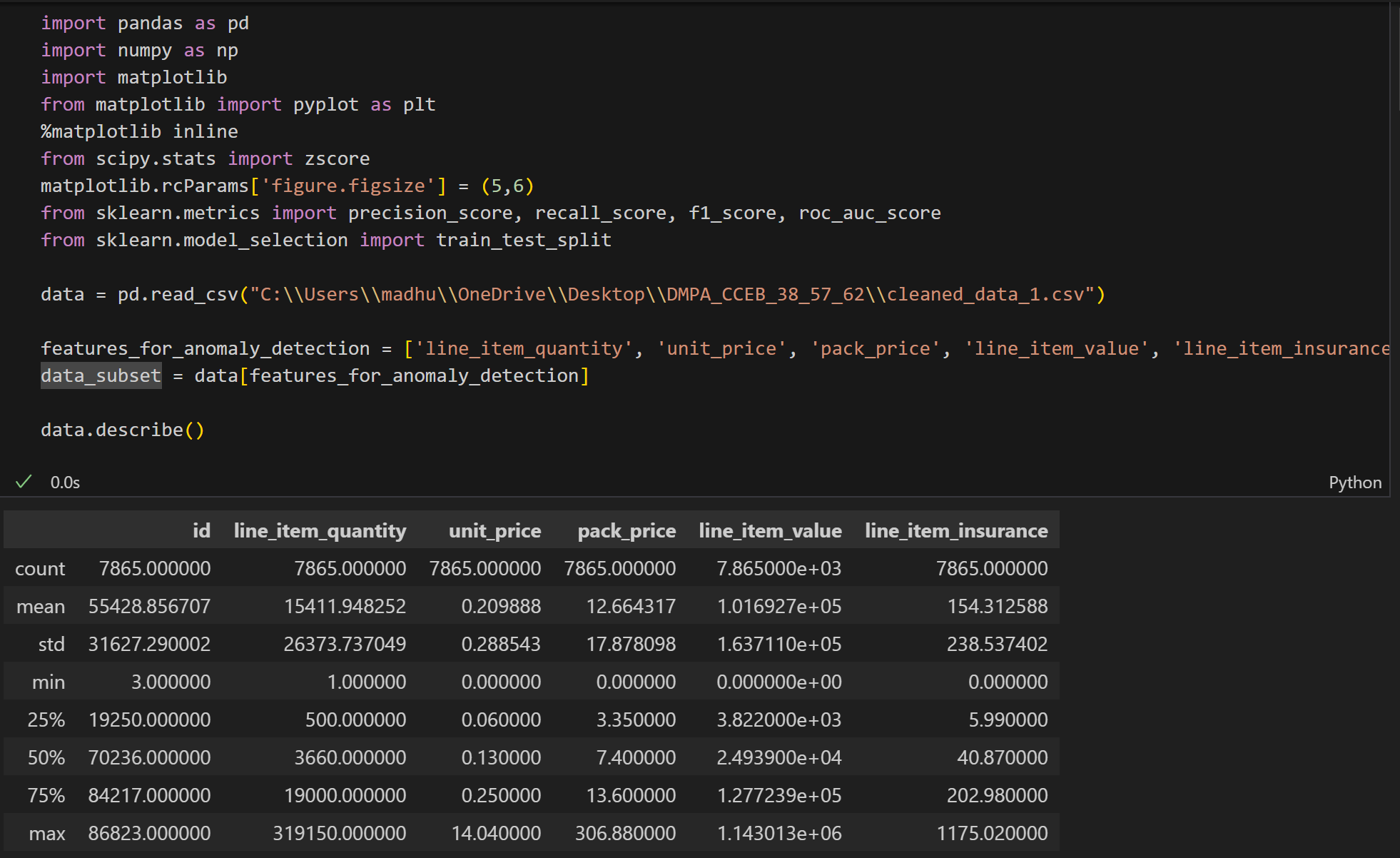
**3. Z-SCORE**

The Z-score algorithm is a statistical technique for identifying outliers or anomalies in a data set. This algorithm is based on the concept of standard scores, also known as Z-scores, which define how many standard deviations a data point has from the mean Z-score of as given data point.

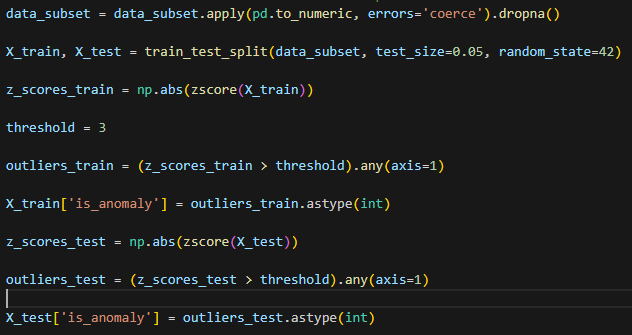
Z-score is calculated using the formula: Z = (X – µ) / σ

The mean is the average value of a data set, and the standard deviation measures the spread or variability of the data. High absolute Z-scores indicate that the data points deviate significantly from the mean. In supply chain management, discrepancies can be associated with unusual or unexpected events, such as sudden increases or decreases in quantities, unit prices or incorrect pack prices.

Step 1: Import the required libraries, read the CSV file, define the list of features (features\_for\_anomaly\_detection) to be used for anomaly detection. Include necessary features like 'line\_item\_quantity', 'unit\_price', 'pack\_price', 'line\_item\_value', 'line\_item\_insurance', and 'weight'. Finding central tendency and data spread using data.describe() method.

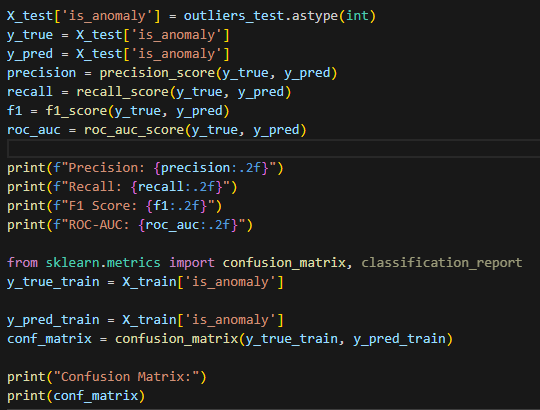


Step 2: Change the data\_subset to numeric, constrain it to NaN and handle the error by discarding the NaN values. Divide the data into training and testing. Calculate the Z-score for the training and testing. Set the Z-score threshold for anomaly detection. Identify outliers in the training set based on the Z-score threshold and create a new column 'is\_anomaly' in the training set to record the outliers



Step-3

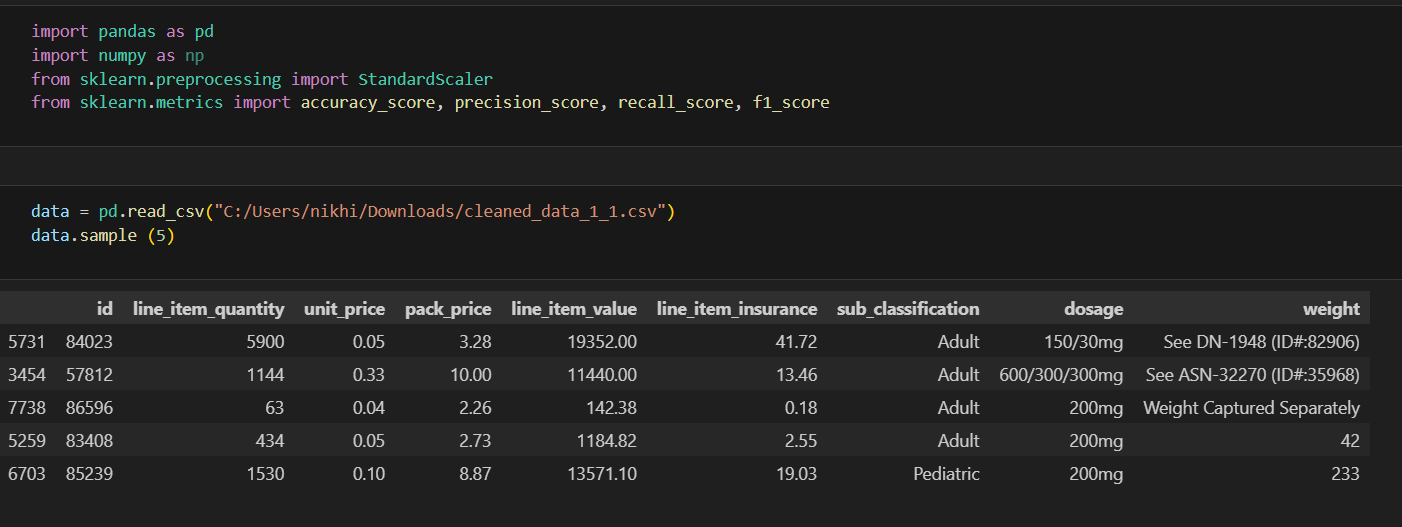
Label outliers in the test set and remove true labels (y\_true) and predicted labels (y\_pred) from the test set. Calculate and display confusion metrics and classification metrics such as precision, recall, F1 score, ROC-AUC



**4)one class support vector machine:**

A one class support vector machine is a statistical technique used to identify outliers or anomalies in the dataset. In one class svm the algorithm is trained on the normal or regular data instances or inlier instances. This algorithm trains on normal data and learns it, then it identifies instances that deviate from normal data and marks them as outliers or anomalies. In One class svm the ‘nu’ value, we can change the strictness of anomalies, higher the nu value higher strictness for detecting anomalies, so a greater number of anomalies and vice versa.

Step 1: Import the required libraries, read the CSV file, print sample data.



Step 2: A function to display the total price.

A screenshot of a computer

Description automatically generated

Step 3: define the list of features (features\_for\_anomaly\_detection) to be used for anomaly detection. Include necessary features like 'line\_item\_quantity', 'unit\_price', 'pack\_price', 'line\_item\_value', 'line\_item\_insurance', and 'weight'. Converting nonnumeric values to NaN values and dropping them and scaling the data.

A black screen with colorful text

Description automatically generated

Step 4: Train the OneClassSvm model using the scaled data and predict the anomalies, then convert predictions to anomalies 1 and normal as 0.

A screen shot of a computer program

Description automatically generated

Step 5: Generation of synthetic labels as the dataset is unsupervised, then initialize all the instances to 0 considering then as normal data and 1 as predicted anomaly, and then evaluating the model by calculating accuracy, precision, recall and f1 score.

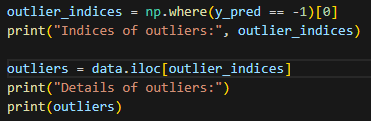
A screen shot of a computer code

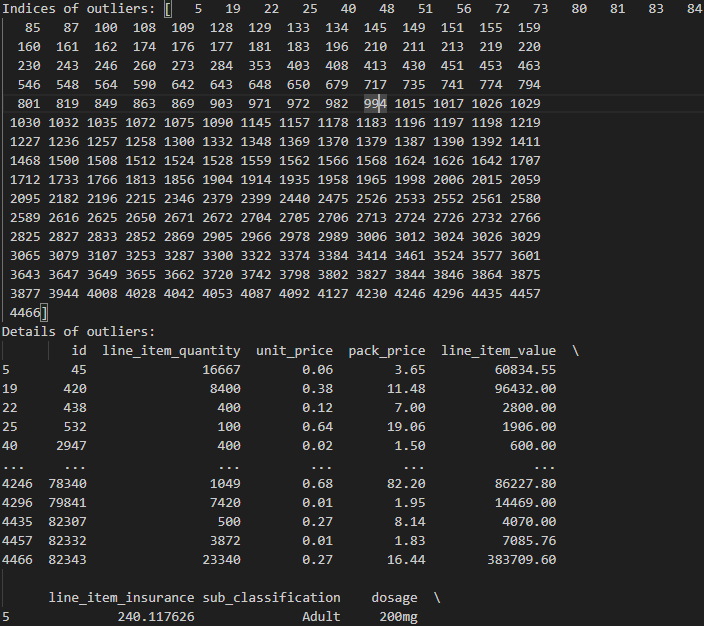
Description automatically generated

**Chapter-4**

**Results & discussions**

**Local Outlier Factor (LOF):**



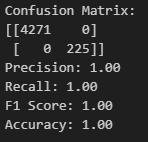


A screenshot of a computer

Description automatically generated

This shows how many anomalies detected based on our lof model.

**Evaluating results of lof:**



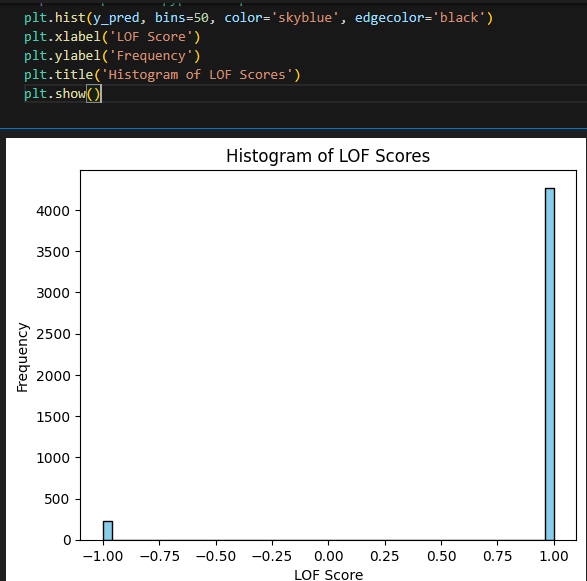
It shows that our model is performing good based on confusion matrix.

This confusion matrix indicates that the dataset having 4271 true negative values which results in that they are correctly predicted as normal class (class 0)

There are 225 true positive values which means that they are detected as anomalies (class 1)

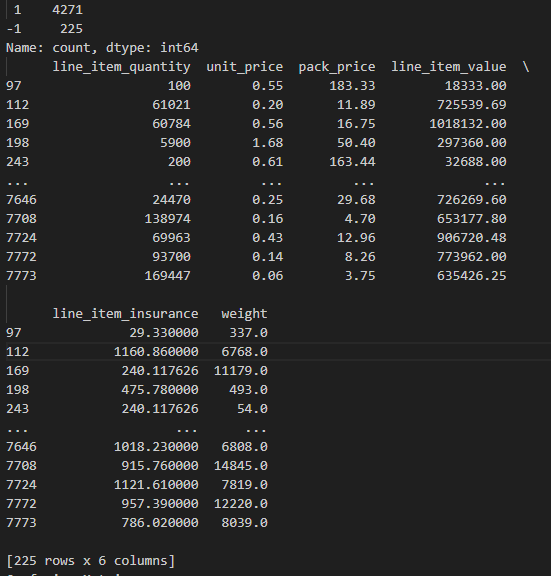
So, after analysing this we can say there are no false positive or false negative values and the precision, recall, f1 score is 1. Since there are no false negative and false positive values the accuracy is 1 which means 100%.

**Visualization:**

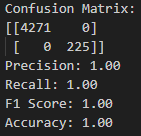


**Results of Isolation Forest:**

**Predicted anomalies:**



**Evaluation of isolation forest model:**



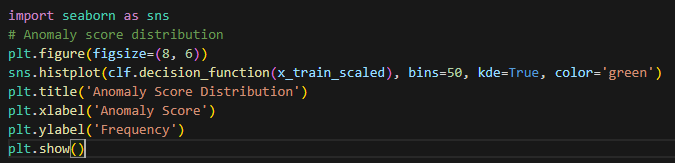
This confusion matrix shows that our code is working very well and finding the anomalies correctly.

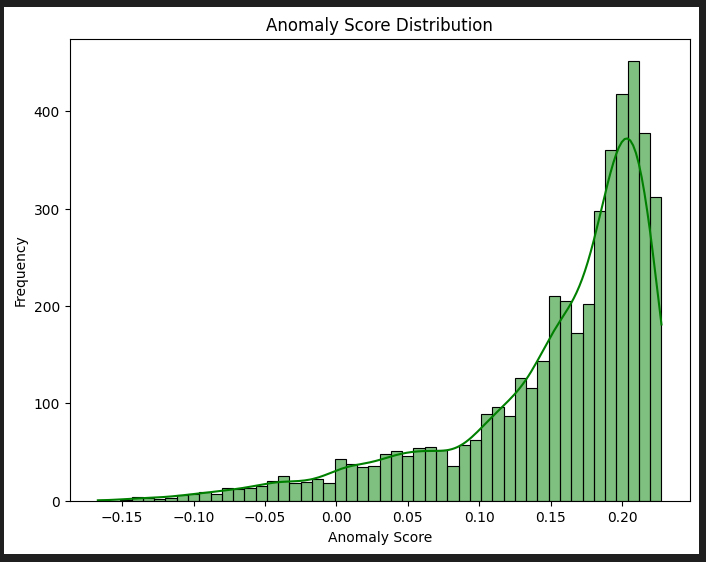
This confusion matrix indicates that the dataset having 4271 true negative values which means they are correctly predicted as normal class (class 0)

There are 225 true positive values which means they are detected as anomalies (class 1)

So, after analysing this we can say there are no false positive or false negative values and the precision, recall, f1 score is 1.

**Visualization:**





**Z-Score:**

**Evaluation of z-score algorithm:**

**A screen shot of a computer

Description automatically generated**

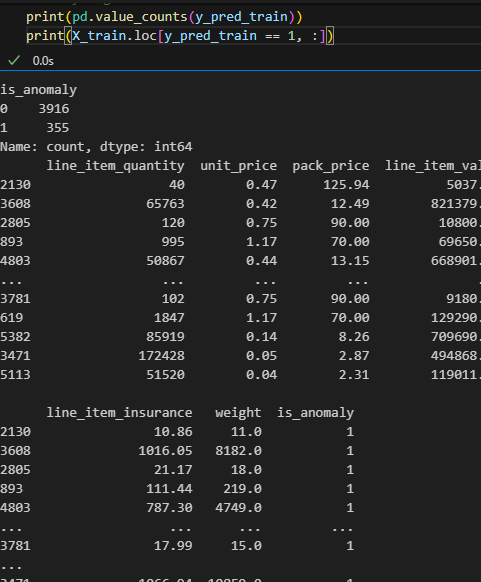
This confusion matrix shows that our code is working very well and finding the anomalies correctly.

This confusion matrix indicates that the dataset having 3916 true negative values which means they are correctly predicted as normal class (class 0)

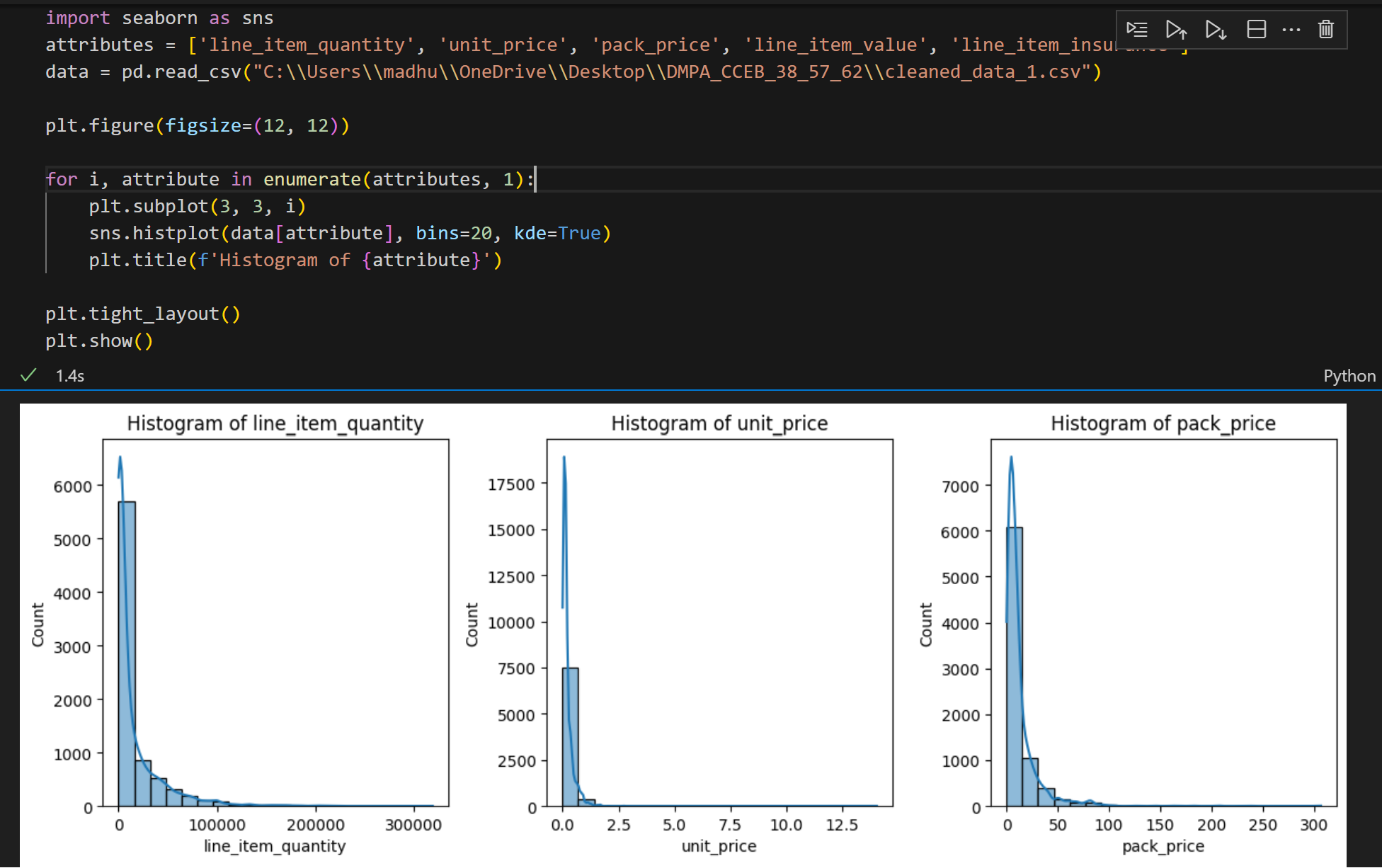
There are 355 true positive values which means they are detected as anomalies (class 1)

So, after analysing this we can say there are no false positive or false negative values and the precision, recall, f1 score is 1.

**Printing the detected outliers and anomalies:**

****

**Data visualization of selected attributes.**



A graph of value and value

Description automatically generated

A screenshot of a computer screen

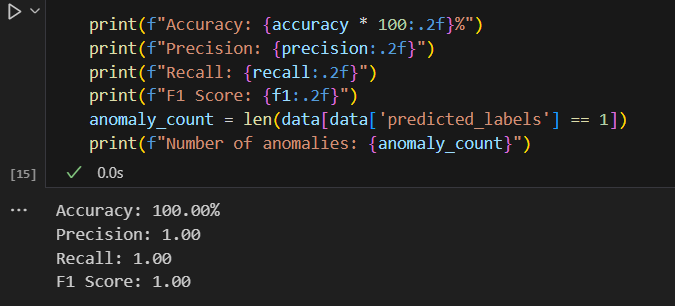
Description automatically generated

A collage of red and blue dots

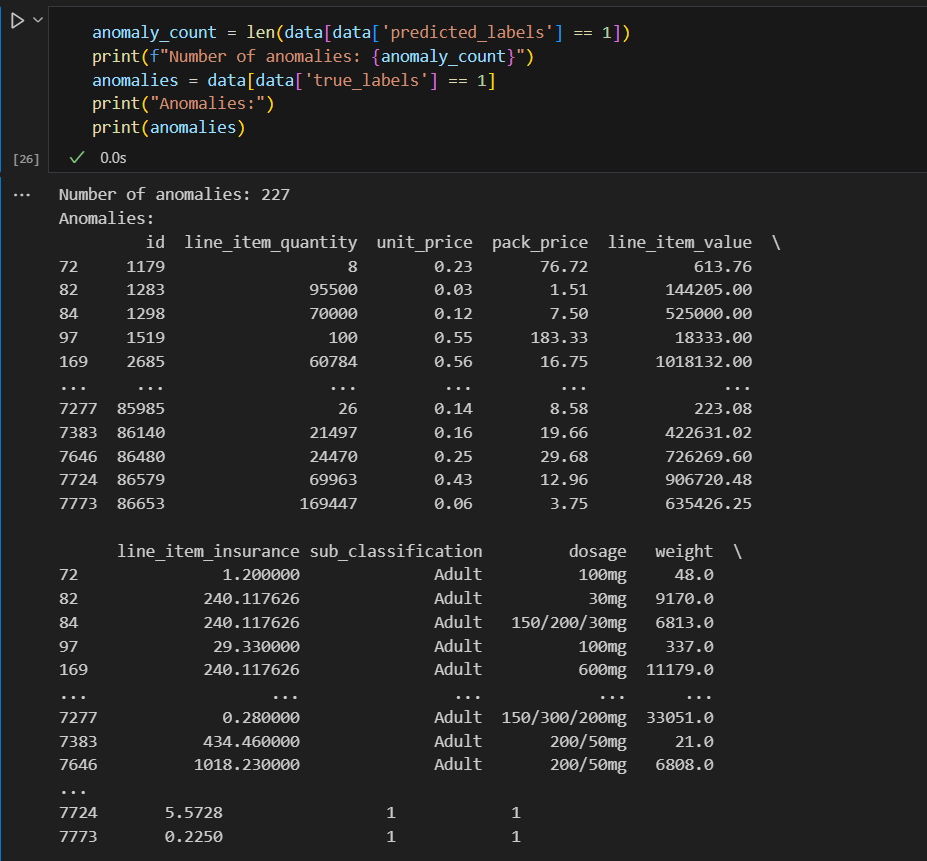
Description automatically generated

**One class svm:**

print the accuracy, precision, recall and f1 score and print the count of anomalies in the dataset considering anomalies are labelled as 1.



Code to count the number of anomalies in the dataset as per nu value and print them.



Visualization:

Data visualization of selected attributes, red are anomalies and blue are normal data. Scatter plot, time series plot and 3d scatter plot.

Timeseries

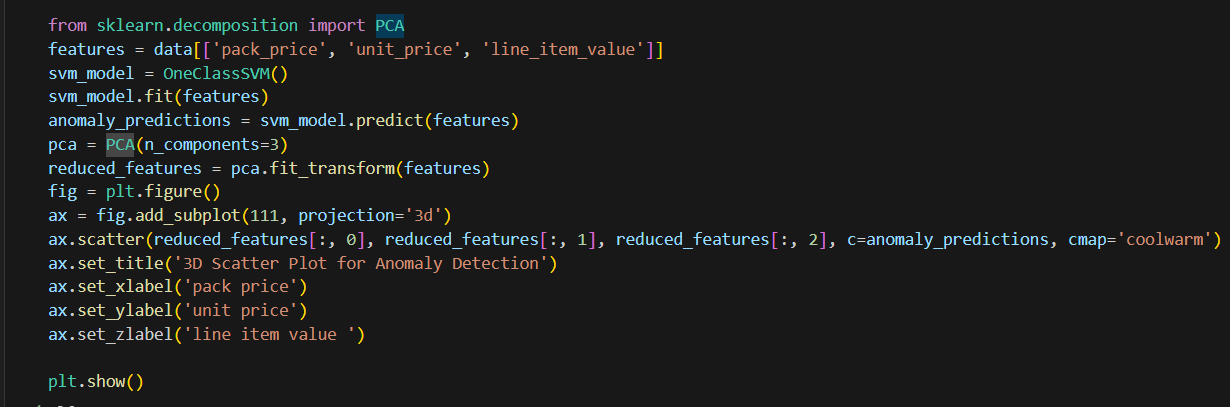
A black background with colorful text

Description automatically generated

A graph showing a number of blue and red dots

Description automatically generated

3d scatter plot



A graph of a scatter plot

Description automatically generated

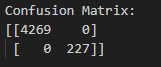
**Confusion matrices obtained:**

Local outlier factor: Isolation Forest:

 A black background with white text

Description automatically generated

z-score: one class svm:

**Anomalies detected:**

**Lof:**

|  |  |  |
| --- | --- | --- |
| **n\_neighbours** | **Contamination** | **Anomalies detected** |
| **15** | **0.05** | **225** |
| **15** | **0.10** | **450** |
| **15** | **0.15** | **675** |
| **15** | **0.20** | **899** |

**Isolation forest:**

|  |  |  |  |
| --- | --- | --- | --- |
| **n\_jobs** | **Random state** | **Contamination** | **Anomalies detected** |
| **15** | **42** | **0.05** | **225** |
| **15** | **42** | **0.10** | **449** |
| **15** | **42** | **0.15** | **675** |
| **15** | **42** | **0.20** | **898** |

**z-score:**

|  |  |
| --- | --- |
| **Threshold** | **Anomalies detected** |
| **2** | **669** |
| **2.5** | **520** |
| **3** | **355** |

**one class svm:**

|  |  |
| --- | --- |
| **Nu** | **Anomalies detected** |
| **0.05** | **227** |
| **0.10** | **448** |
| **0.15** | **674** |
| **0.20** | **899** |

**Lof:**

**pros**: LOF is effective in detecting local outliers and is generally robust. It provides a measure of the density of data points in the area.

**Cons:** The detected anomalies seem to increase linearly with the contamination level.

**Isolated forest:**

**Pros:** Isolation Forest is more efficient and scalable. Generally effective in detecting anomalies in high-dimensional data.

**Cons**: As with LOF, detected abnormalities increase with contamination.

**Z-Score:**

**Pros**: The Z-Score is simple and interpretable. It measures how many standard deviations a data point is from the mean.

**Cons:** As the threshold increases, the detected anomalies decrease, which may indicate that the method is sensitive to the choice of threshold and may also assume normality in data distribution.

**One-class SVM:**

**pros**: One-class SVM is effective in capturing the underlying structure of all available data.

**Cons**: Similar to LOF and Isolation Forest, the number of detected anomalies increases with the fraction of the parameter (nu) representing the upper limit of margin errors

**Chapter 5**

**Conclusion:**

Supply chain anomaly detection system provides a simple and valuable solution for identifying irregularities in supply chain data.

Through outlier analysis techniques discussed above the system classifies and analyses features in the data set, providing insight into potential anomalies. The anomaly detection system provides decision support for stakeholders, including supply chain managers. By updating the data used for training, the system can learn from new systems and adapt its anomaly detection capability to changing conditions.

One-Class SVM is a good candidate if we have aspirations for an algorithm that considers the underlying structure of all our data.

If performance and scalability are important, and we have high quality data, an isolated forest may be the right choice.

Finally, One class svm is detecting highest anomalies so it is the best algorithm among the four.

**References:**

**Dataset from: data.gov.in**

[Supply Chain Shipment Pricing Data - Catalog](https://catalog.data.gov/dataset/supply-chain-shipment-pricing-data-07d29)

**[1]** Baumgartner, P., & Krumpholz, A. (2021). Anomaly Detection in a Boxed Beef Supply Chain. ACM International Conference Proceeding Series, 1–7. <https://doi.org/10.1145/3474963.3474964>

**[2]** Jiménez-Carmelo, A. M., Li, P., Erasmus, S. W., Wang, H., & van Ruth, S. M. (2023). Spatial-Temporal Event Analysis as a Prospective Approach for Signalling Emerging Food Fraud-Related Anomalies in Supply Chains. Foods, 12(1). <https://doi.org/10.3390/foods12010061>

**[3]** Sahin, B., & Soylu, A. (2020). Multi-Layer, Multi-Segment Iterative Optimization for Maritime Supply Chain Operations in a Dynamic Fuzzy Environment. IEEE Access, 8, 144993–145005. <https://doi.org/10.1109/ACCESS.2020.3014968>

**[4]** Yeboah-Ofori, A., Islam, S., & Brimicombe, A. (2019). Detecting cyber supply chain attacks on cyber physical systems using bayesian belief network. Proceedings - 2019 International Conference on Cyber Security and Internet of Things, ICSIoT 2019, 37–42. <https://doi.org/10.1109/ICSIoT47925.2019.00014>

**[5]** Chukkapalli, S. S. L., Ranade, P., Mittal, S., & Joshi, A. (2021). A Privacy Preserving Anomaly Detection Framework for Cooperative Smart Farming Ecosystem. Proceedings - 2021 3rd IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications, TPS-ISA 2021, 340–347. <https://doi.org/10.1109/TPSISA52974.2021.00037>

**[6]** Dimitrios Tychalas, Anastasis Keliris, Michail Maniatakos, “Supply Chain threats for stealthy data exfiltration in industrial control systems”, “IEEE”, 2019. <https://ieeexplore.ieee.org/abstract/document/8854451/>

**[7]** Zunic, E., Tucakovic, Z., Delalic, S., Hasic, H., & Hodzic, K. (2020). Innovative Multi-Step Anomaly Detection Algorithm with Real-World Implementation: Case Study in Supply Chain Management. 2020 IEEE / ITU International Conference on Artificial Intelligence for Good, AI4G 2020, 9–16. <https://doi.org/10.1109/AI4G50087.2020.9311045>

**[8]** Vu, D. L., Pashchenko, I., Massacci, F., Plate, H., & Sabetta, A. (2020). Towards Using Source Code Repositories to Identify Software Supply Chain Attacks. Proceedings of the ACM Conference on Computer and Communications Security,

2093–2095. <https://doi.org/10.1145/3372297.3420015>

**[9]** Rejeb, A., Keogh, J. G., & Treiblmaier, H. (2019). Leveraging the Internet of Things and blockchain technology in Supply Chain Management. Future Internet, 11(7). <https://doi.org/10.3390/fi11070161>

**[10]** Tsoukas, V., Gkogkidis, A., Kampa, A., Spathoulas, G., & Kakarountas, A. (2022). Enhancing Food Supply Chain Security using Blockchain and TinyML. Information (Switzerland), 13(5). <https://doi.org/10.3390/info13050213>