Anomaly Detection of IoT Data for Hemodialysis

May 2024

Lana al dossary

Introduction

Discovering sequences or occurrences in a dataset that substantially differ from expected or normal behavior is known as anomaly detection. This process can be quite helpful in guaranteeing the efficacy and safety of hemodialysis treatment for patients. The idea also intends to create a system for detecting anomalies in loT data produced by hemodialysis . Healthcare providers can identify potential issues early and take appropriate action to prevent complications by looking for anomalous patterns in the data.

Problem Statement

Building an accurate anomaly detection system that can spot anomalous trends in loT data taken from hemodialysis datasets is the goal. Using a dataset of different sensor readings and operational parameters from loT devices, our goal is to construct an entity relationship model that can effectively discriminate between typical and anomalous behavior.

Sample Data

We provide an example dataset containing sensor for hemodialysis devices:

Patient Information	Treatment Information	Heart rate	Temperature
○ Patient ID : 11231874	○ Date and time of treatment: 2/01/2024	67 beats per minute	37 C
○ Patient ID: 11432985	○ Date and time of treatment: ○ 17/01/2024	70 beats per minute	36.8 C
○ Patient ID: ○ 19872165	○ Date and time of treatment: ○ 27/02/2024	74 beats per minute	36.5 C

Methodology

Participants will develop an anomaly detection system using different techniques to analyze the IoT data from hemodialysis.

1. **Data Collection**: Establish information from numerous sensors and IoT devices that are utilized during hemodialysis treatments. This information can be obtained from the National Data Platform, as well as through administering a survey to gather more data. It may include the patient's vital signs, machine performance parameters, fluid parameters, and other pertinent information.

2. **Model Selection:** Based on your particular use case, select the best anomaly detection algorithm. Real-time data can be found through a variety of technology and technique related to information correlation.

3.**Real-time Anomaly Detection:** With fresh, untrained data gathered from hemodialysis sessions, apply the trained model. Keep an eye on the incoming IoT data and mark as potentially anomalous any instances that surpass the specified threshold. Request involvement and preventive measures are made possible by real-time detection. by employing signature-based forecasting techniques, which employ blacklists to define sets of patterns representing known threats or undesired behaviors and require manual specification based on historical data.

This methodology allows participants to focus on developing effective anomaly detection systems for IoT data from home hemodialysis devices while facilitating evaluation of their models' performance on unseen data during the evaluation phase. It encourages participants to explore various techniques and optimize their models for accurate detection of anomalies, ensuring the safety and reliability of home hemodialysis treatment.

Code

A sample using Python code for anomaly detecting anomalies in high-dimensional datasets:

Database.py

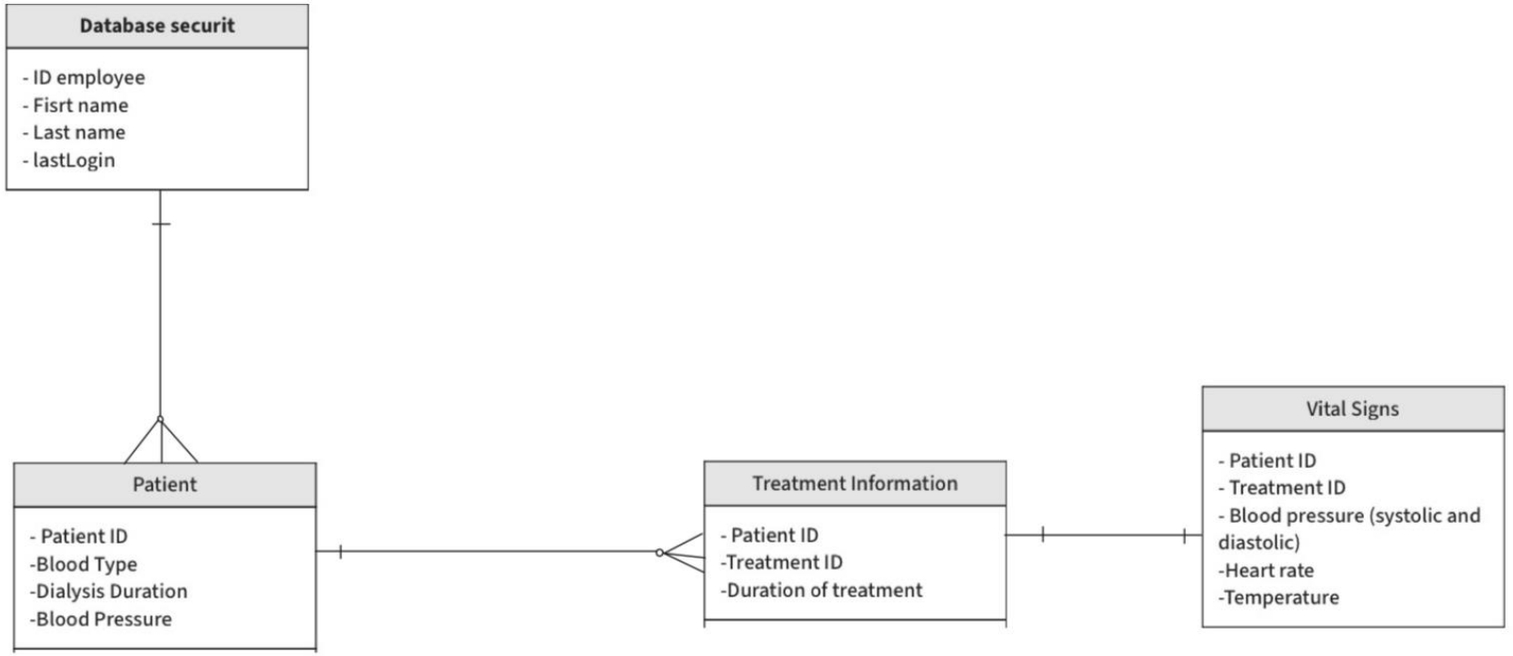
```python
import numpy as np
from sklearn.ensemble import IsolationForest

# Generate a high-dimensional dataset for demonstration
np.random.seed(50)
n_samples = 500
n_features = 10
X = np.random.randn(n_samples, n_features)

# Create an Isolation Forest instance
isolation_forest = IsolationForest(contamination=0.05, random_state=50)

# Fit the model to the data
isolation_forest.fit(X)

# Predict anomaly scores for each data point
anomaly_scores = isolation_forest.decision_function(X)

# Determine the outliers/anomalies
outliers = np.where(anomaly_scores < 0)[0]

# Print the indices of the detected anomalies
print("Detected anomalies:")
print(outliers)
```

Sample Outputs for the code :

After creating model, we obtained the following output

Detected anomalies :[1 5 10 25 29 32 34 43 48 53 55 56 67 81
84 86 95 97 99 101 102 114 116 118 126 131 134 135 137 146 151 156
158 160 162 163 165 176 178 181 182 183 189 191 193 194 195 200 209
213 216 218 219 221 228 230 232 237 239 241 248 255 257 263 266 267
270 276 280 281 282 284 287 297 299 301 303 306 307 312 318 324 325
327 328 329 331 332 335 337 339 344 345 346 349 350 351 353 357 358
359 364 366 368 369 371 372 375 377 378 383 389 390 392 394 395 397
398 400 405 407 411 413 415 416 417 418 420 422 423 425 426 428 430
431 433 435 437 438 442 448 450 451 453 454 456 457 460 462 464 465
466 470 471 472 473 474 475 476 477 481 482 484 487 488 490 493 494
495 497]

Entity relationship diagram



Conclusion

The purpose of database security is to find and correct security flaws that could have the availability, confidentiality, and integrity of data. Anomaly detection, which combines data analysis methods with security controls to find odd or suspicious activity in databases, is a significant component of this research. Organizations can monitor database activities and identify deviations that might point to security threats or breaches by setting up a baseline of normal behavior. Analyzing database logs and applying machine learning algorithms to spot anomalies are common methods. Strong security measures like access controls, encryption, and audits are required to guarantee long-term data viability. Organizations can preserve real-time data and guarantee its long-term security by proactively detecting and addressing security issues through the integration of data analysis and security measures.