VERITAS illuminate 2024

uConnect Hackathon 2024

IntelliDetect

Team Zephyr

Radhika Fadnavis

Srushti Johari

Gayatri Joshi

MKSSS's Cummins College of Engineering, Pune

Business Case Slide

Idea Description	
Problem Statement	Anomaly detection using data access patterns Write Anomaly detection for Windows/Linux Unstructured file data or NAS file server that analyses unusual user activity and user behavior. User behavior is represented as any user actions performed on the system. Consider using capabilities of File Change Log, API usage, Audit logs, WORM, CPU usage, and unusual disk activity. Leverage AI/ML techniques. Understand different attack patterns and resemble to actions carried out. The algorithm should demonstrate accuracy and consider false positives and false negatives.
Solution Benefits	Early detection of unauthorized access or abnormal user behavior. Improved security posture by identifying potential threats before they cause damage. Reduced risk of data breaches and data loss. Increased operational efficiency by automating the detection of anomalies.
Category	AI/ML
Theme	Cybersecurity and AI/ ML

The anomaly detection system on unstructured file data

The application captures real time parameters from the machine which are fed to the models that recognises whether there exists an anomaly.

The user will be given a warning through a pop up message to notify the anomalous system behavior.

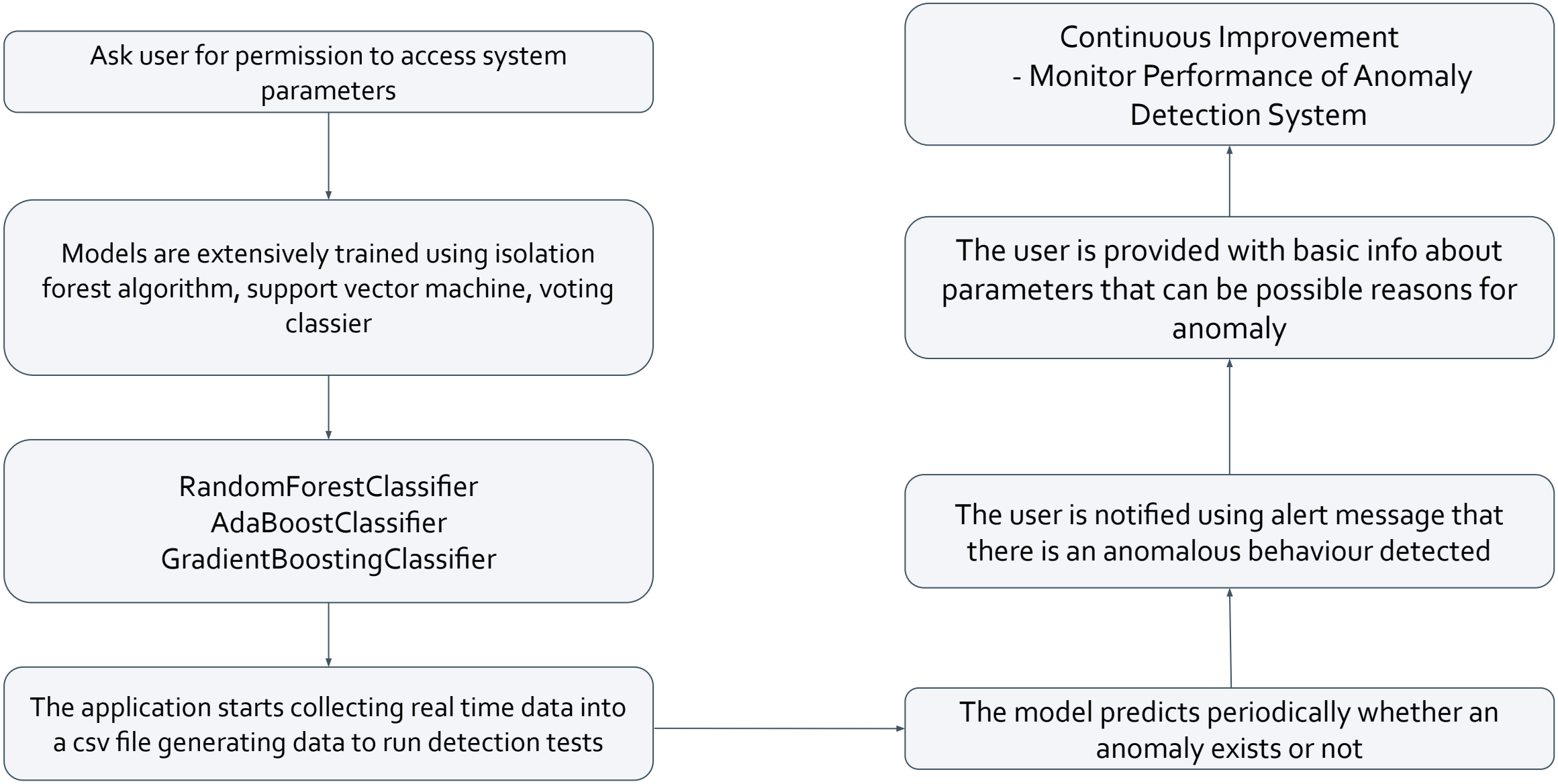
Challenges we had:

1. Unavailability of labeled dataset.

   This led to working on unsupervised models.

2. False alarm rate due to scarce and varied data from multiple machine systems having different configurations.
3. Access to some files are restricted due to permission and privileges.

Application Flow



Parameters considered for model prediction

NETWORK PARAMETERS

1. **Login** parameters like user and guest session entries that validates authenticity
2. **Bytes sent** and **received** through the network
3. **Destination host count**

HARDWARE PARAMETERS

1. Involve **cpu** parameters that deal with percent utilization of resources
2. **Disk** parameters like read bytes and write bytes per second.
    a. This can detect excessive usage of disk space.
3. **File change logs** record metadata like timestamp, file path, user ID, and process ID responsible for the change

Tech stack to be used

- **Isolation Forest algorithm** for outlier detection
- **SVM (Support Vector Machine)**
- **Random Forest:** It aggregate predictions from multiple decision trees, offering high accuracy, robustness against overfitting, and handles both numerical and categorical features using Python.
- For UI: **PyQt5**
- For accessing system parameters: **psutil library**

Business Value

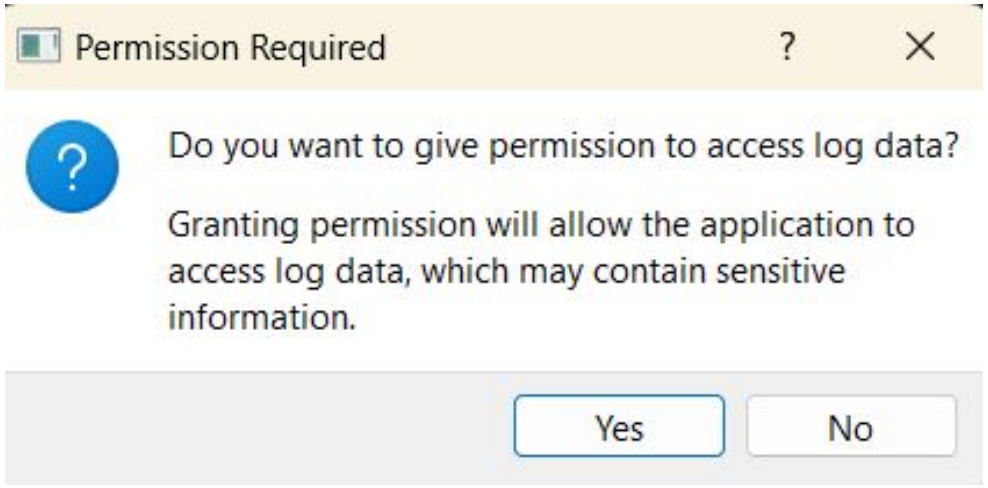
The anomaly detection system works as a desktop application which can be run on a user machine to detect unusual activity in system.

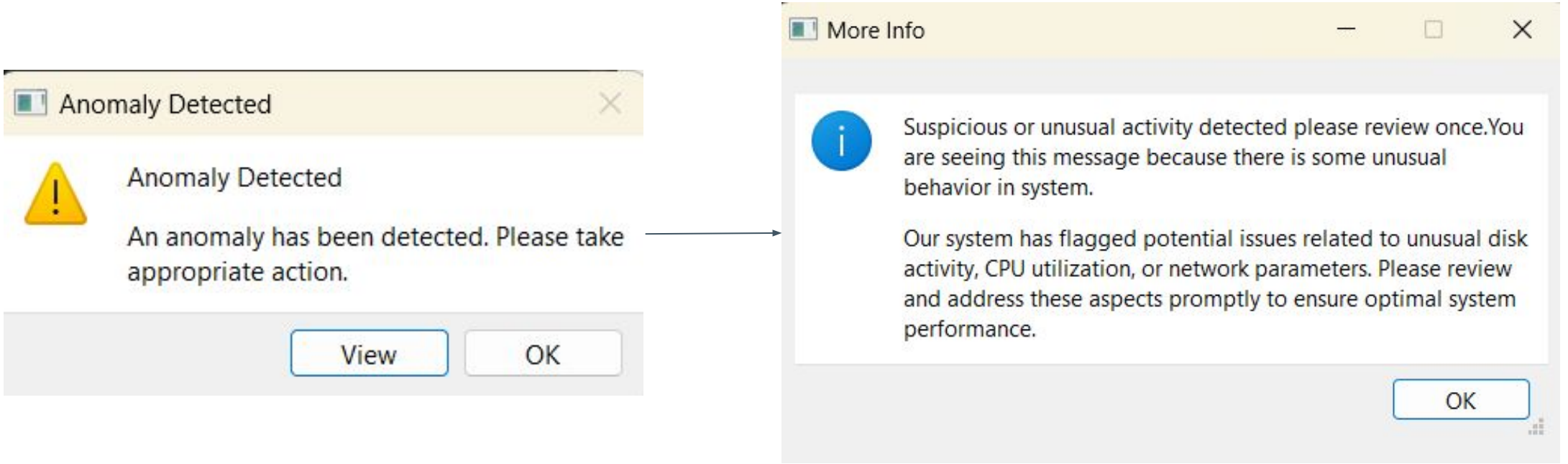
1. The application boasts a user-friendly design with a straightforward and easily navigable interface, making it accessible to users of all levels of expertise.
2. The application is designed to be highly scalable, accommodating increased demand and evolving user requirements.
3. The application is lightweight, ensuring seamless performance without imposing a heavy load on system resources.
4. Compatible with all Windows machines and configurations, the application ensures a universal user experience across diverse systems."

Time to market and Innovative & Efficient

The product can be swiftly introduced to the market, potentially within a few months, as a system application compatible with Windows machines.
It boasts *efficiency* through **multithreading**, ensuring smooth operation.
User-friendly in design, the application supports **continuous improvement** through ongoing model training data updates and *meticulous data monitoring* processes.

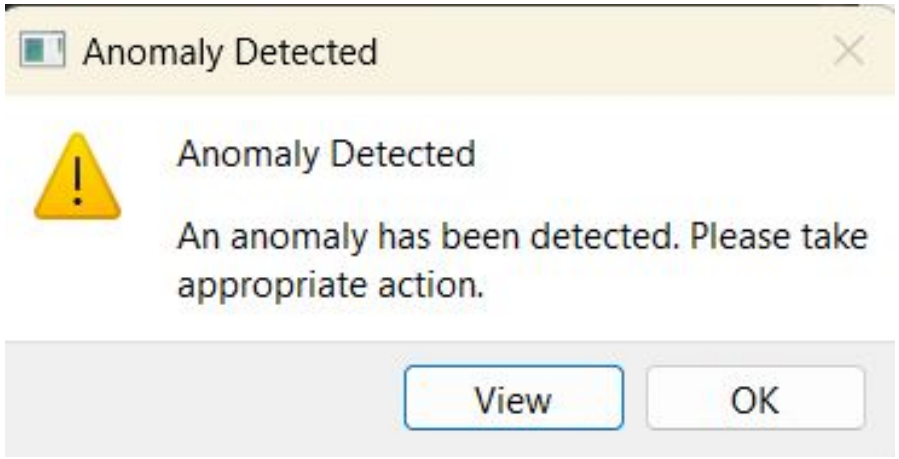
UI elements

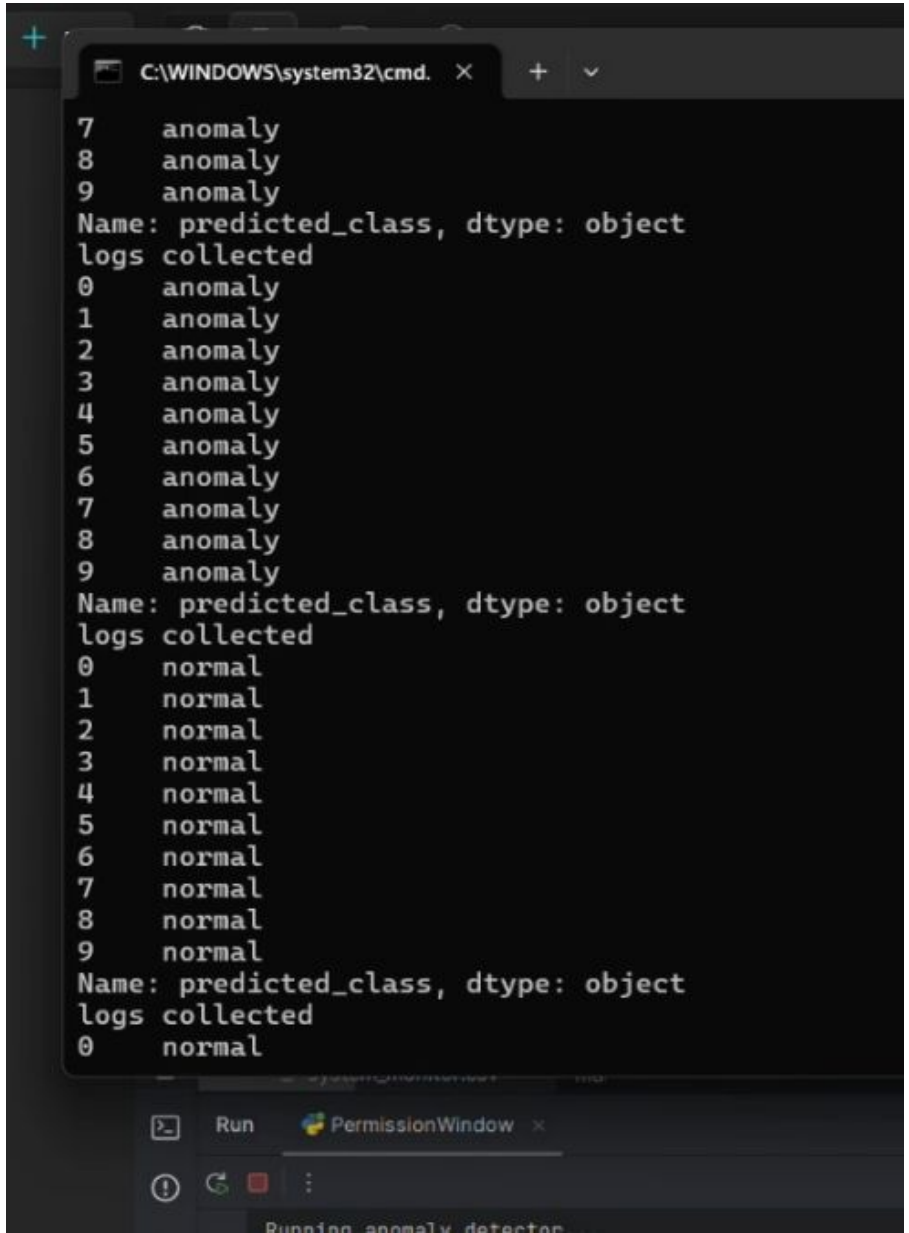

Permission Required
Do you want to give permission to access log data?
Granting permission will allow the application to access log data, which may contain sensitive information.
Yes
No


Anomaly Detected
Anomaly Detected
An anomaly has been detected. Please take appropriate action.
View
OK
More Info
Suspicious or unusual activity detected please review once.You are seeing this message because there is some unusual behavior in system.
Our system has flagged potential issues related to unusual disk activity, CPU utilization, or network parameters. Please review and address these aspects promptly to ensure optimal system performance.
OK

Successfully detected real-time anomaly during testing

During a testing phase, a sudden network lag was encountered. The application adeptly identified this anomalous network behavior and promptly issued a warning alert on the monitor. Hence we can clearly state that the application works well in detecting unusual behaviour.





X-factor

- **Complete Real Time Application**- The application collects actual parameters of the system and feeds to the model. The user can monitor anomalous behavior of system real time.

- **Constant Improvement** The application uses supervised modeling for network, This is constantly updating the training data to refine the model gradually. After generating an output, the collected data and prediction is added to the training set. This iterative approach ensures the ongoing improvement of our model.

- **Multithreading** for multiple programs to run simultaneously. Detection for network and hardware (cpu, disk work in parallel).

  Since the program is running continuously in the background, detection for each should be continuously done in parallel. ->Lightweight application

Future Scope

1. Scalability to Cloud Resources: for efficient storage of large datasets.

2. Enhanced Model Parameters and Accuracy:
With an increased number of parameters, the model achieves improved accuracy in its predictions.

3. Quality and Quantity of Data Improvement: Continuous efforts on the training data, ensures a more robust model.

4. Feature to view anomaly statistics: the capability to view detailed information about the source of anomalies and other info.

5. Automated AI-Driven Incident Response: The application incorporates automating responses to identified anomalies.

6. Utilization of Daemon Processes:
To ensure smooth background operation, the system employs daemon processes, enhancing overall efficiency.

Team Photo

Radhika Fadnavis

Srushti Johari

Gayatri Joshi

VERITAS

illuminate 2024

Thank you