AUTOMATED RANSOMWARE DEFENSE SYSTEM USING BILSTM AND DATA PROTECTION

V. Anandham Nithiya

Department of Computer Science, St. Mary's College (Autonomous) Thoothukudi

Affiliated to Manonmaniam Sundaranar University, Tirunelveli, Tamilnadu, India.

ABSTRACT

Ransomware attacks pose a severe cybersecurity threat, encrypting critical files and demanding payment for decryption, often without guaranteeing recovery. Traditional detection methods struggle against evolving ransomware variants, leaving users and organizations vulnerable. The Project Automated Ransomware Defense System Using BiLSTM and Data Protection introduces an advanced, autonomous solution to combat cryptographic ransomware. Utilizing a BiLSTM (Bidirectional Long Short-Term Memory) model**,** the system analyzes temporal behavior patterns to detect and block ransomware in real time, ensuring adaptive and robust protection. In addition to detection, the system employs Format Preserving Encryption (FPE) to proactively safeguard data by camouflaging files with excluded extensions and hiding them in obscure directories, minimizing the risk of unauthorized encryption. Unlike traditional signature-based defenses, this system integrates dynamic behavioral analysis and proactive data protection, enhancing resilience against both known and emerging threats. By autonomously mitigating ransomware attacks, this system provides a comprehensive and scalable defense for individuals, businesses, and government entities, securing critical systems and sensitive data.

KEYWORDS – *System Security, Algorithm, BiLSTM, FPE, File Guardian*

1. INTRODUCTION

Ransomware is a rapidly evolving cybersecurity threat that encrypts user or organizational files and demands payment for decryption. Over the past decade, Ransomware has become a significant cyber threat to both individuals and organizations[1]. High-profile incidents like WannaCry, Snake, and Pemex have demonstrated the devastating impact of ransomware, resulting in significant financial losses, service disruptions, and data breaches. Despite continuous advancements in cybersecurity, traditional detection and prevention mechanisms struggle to combat the ever-changing nature of ransomware, particularly zero-day exploits and novel variants. Existing detection methods, such as signature-based approaches, are often ineffective

against new and unknown threats. While machine learning-based techniques have shown promise, they face challenges such as dataset imbalances, high false positive and false negative rates, and substantial computational demands. Many conventional solutions also lack adaptability, scalability, and proactive threat mitigation capabilities, leaving systems vulnerable to advanced ransomware attacks.

To address these limitations, The Project aims to develop an Automated Ransomware Defense System that enhances detection accuracy and strengthens data protection against evolving ransomware threats. It also introduces a proactive defense strategy that integrates BiLSTM (Bidirectional Long Short-Term Memory) and File Guardian technology to improve ransomware detection and mitigate encryption risks. The BiLSTM model dynamically analyzes ransomware behavior patterns, enabling early and accurate threat identification, even for novel variants. In parallel, File Guardian prevents unauthorized encryption by securely managing file access and preserving data integrity. This dual-layered approach ensures both real-time detection and proactive prevention, significantly reducing the impact of ransomware attacks.

The Project focuses on designing a scalable, adaptive, and user-friendly cybersecurity solution that incorporates real-time monitoring, automated alerts, and continuous learning mechanisms to adapt ransomware. By leveraging deep learning and cryptographic techniques, and aims to create a comprehensive, efficient, and resilient security system that safeguards critical data and systems against modern ransomware threats.

2. MATERIALS AND METHODS

2.1. System Design

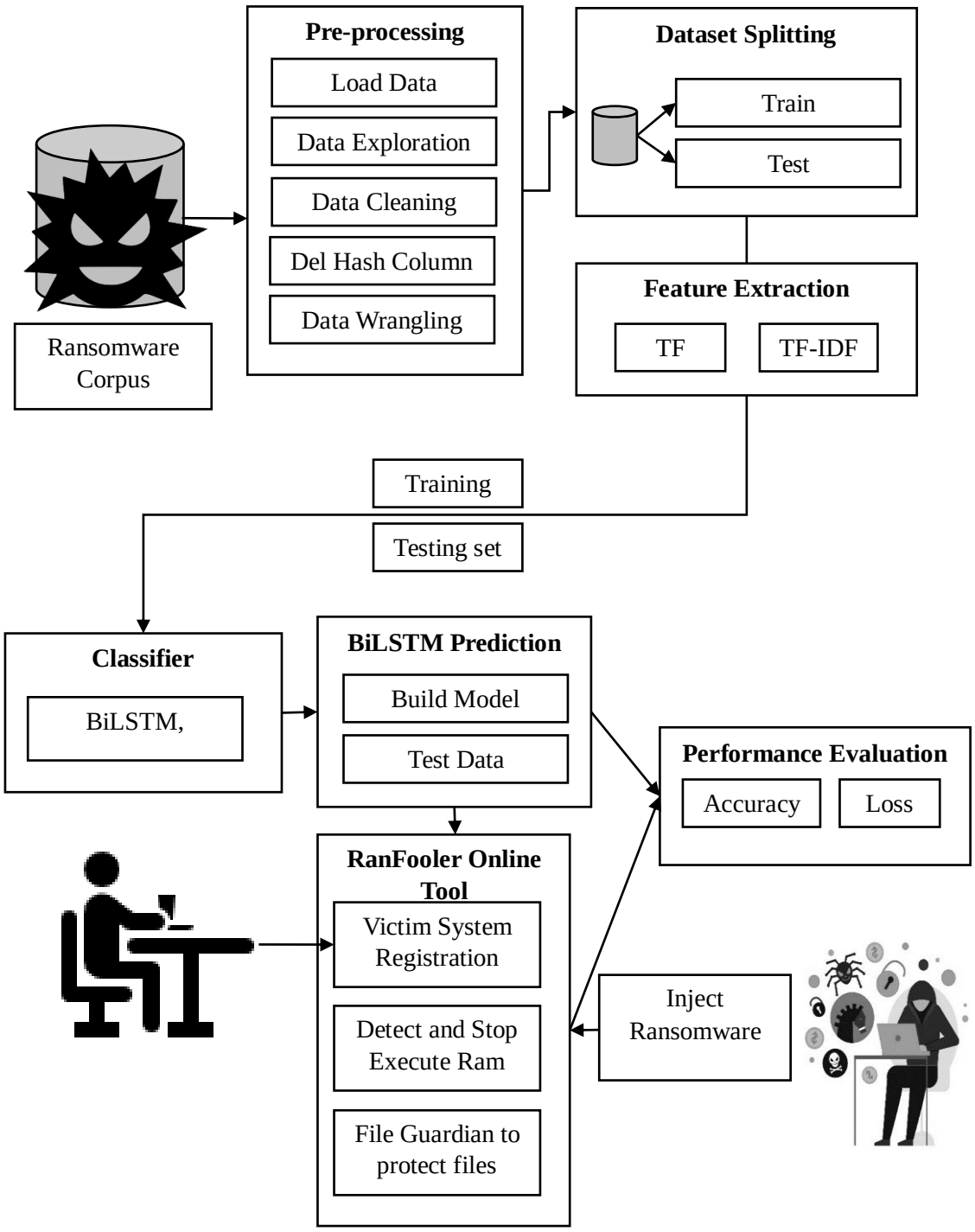

Figure 1: Architecture Design of Ransomware Defense System

2.2. Data Collection

The project gathers network threat intelligence data, including detailed information on ransomware attack incidents, for analysis. and evaluate the ransomware detection system[2]. The dataset consists of various ransomware samples along with benign files, enabling the model to differentiate between malicious and

legitimate activities. It includes Portable Executable (PE) files, which contain detailed assembly instructions used for static and behavioral analysis.

The data was sourced from public cybersecurity repositories, malware databases, and research datasets, ensuring a diverse set of ransomware variants. Preprocessing steps were performed to clean the dataset, extract relevant features, and balance the data to avoid bias in model training. This dataset serves as the foundation for training the BiLSTM model, allowing it to recognize ransomware patterns and improve detection accuracy.

2.3. Data Analysis Techniques

Machine learning (ML) Algorithms used to teach machines how to handle the data more efficiently. The ransomware detection system leverages Bidirectional Long Short-Term Memory (BiLSTM) for behavior analysis and Format Preserving Encryption (FPE) for data protection[3]. These techniques work together to enhance detection accuracy and safeguard critical files from ransomware attacks.

BiLSTM (Bidirectional Long Short-Term Memory)

- A deep learning model that analyzes ransomware behavior in both forward and backward directions.
- Recognizes complex temporal patterns, making it effective in identifying ransomware based on system behavior and file modifications.
- Learns from past and future contexts, improving adaptability to evolving ransomware threats.

FPE (Format Preserving Encryption)

- Encrypts files while maintaining their original format, preventing unauthorized modifications.
- Stores protected files in hidden directories to reduce ransomware exposure.
- Ensures that ransomware cannot modify critical files by disguising them with excluded extensions.

Algorithm: Ransomware Detection Using BiLSTM

1. *Input :* File access logs, system behavior data
2. *Preprocessing:*
    - Normalize File Behavior Data
    - Tokenize system logs into sequential input
3. *Feature Extraction:*

- Extract timestamps, file types, access frequency, and modification patterns.
- Encode Sequential behavior into numerical representations.

4. *BiLSTM Model Training:*
   - Initialize BiLSTM layers with forward and backward processing
   - Train model on labeled ransomware and non-ransomware data
   - Optimize loss function using Adam optimizer
5. *Real Time Detection:*
   - Feed new system logs into trained BiLSTM
   - Generate probability scores for ransomware classification
   - If probability exceeds a threshold → Alert user and trigger file protection mechanism
6. *Output:* Ransomware detection alert and activation of FPE for file protection.

Multi-Layered Defense Strategy:

- *Real-time Ransomware Detection:* BiLSTM continuously monitors user activity and system interactions.
- *Proactive Defense:* FPE automatically encrypts critical files, preventing unauthorized access.
- *Adaptive Learning:* The system updates itself to counter new ransomware variants.
- *Automated Alerts:* Users receive notifications when potential ransomware activity is detected

2.4. Tools and Software Used

- *Operating System – Windows 10:* Used as the primary development and testing environment, providing stability, compatibility with required software, and support for machine learning frameworks.
- *Python:* Python is a general purpose programming language that supports multiple programming paradigms. A high-level programming language that takes very few lines compared to other programming languages to accomplish a task. The language comes with large number of built-in methods as a part of standard library[4].
- *TensorFlow:* An open-source machine learning framework used for building and training deep learning models. It provides extensive tools for neural

networks, making it ideal for implementing the BiLSTM model in ransomware detection. TensorFlow enables efficient computation and model optimization for large-scale data processing.

- *Pandas:* A powerful data analysis and manipulation library that simplifies handling structured data. It provides high-performance data structures and functions for data cleaning, transformation, and statistical analysis, making it essential for preprocessing ransomware datasets.
- *NumPy:* A core library for numerical computing in Python, providing support for large, multidimensional arrays and matrices. It enables efficient mathematical operations and is crucial for handling complex datasets required for deep learning models like BiLSTM.
- *Scikit-learn***:** A machine learning library offering a range of classification, regression, and clustering algorithms. It is used for feature extraction, model evaluation, and performance tuning, complementing the deep learning capabilities of TensorFlow.
- *Flask:* A lightweight and flexible web framework used for developing backend services and APIs. It facilitates the integration of the ransomware detection system with a web-based interface, allowing users to interact with the security system seamlessly.
- *MySQL:* Databases play a crucial role in data management and information technology. They serve as foundational infrastructures storing essential data relied upon by numerous applications and systems [5].

3. Results and Discussion

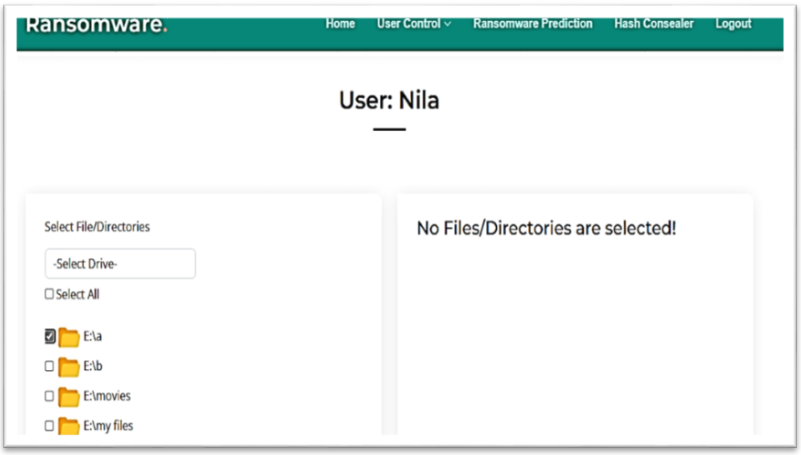

Figure 2: File Selection

Figure 2 shows an interface which allows an authenticated users to select specific files and directories for protection against ransomware attacks.

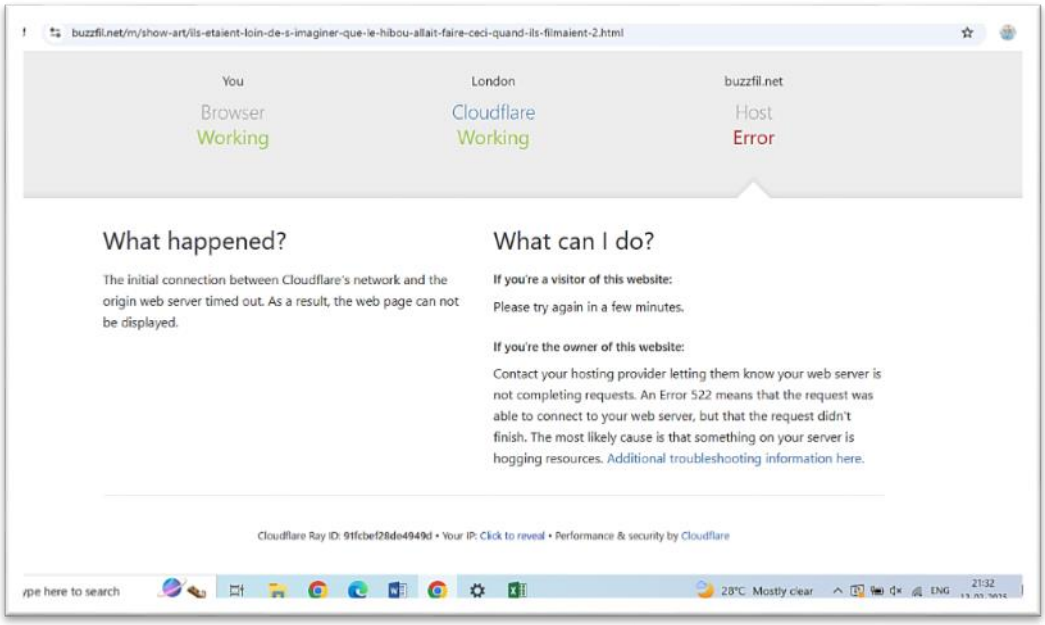

Figure 3: Browsing Malicious Website

In Figure 3 User Accessing a malicious website which can trigger ransomware infections by exploiting system vulnerabilities.

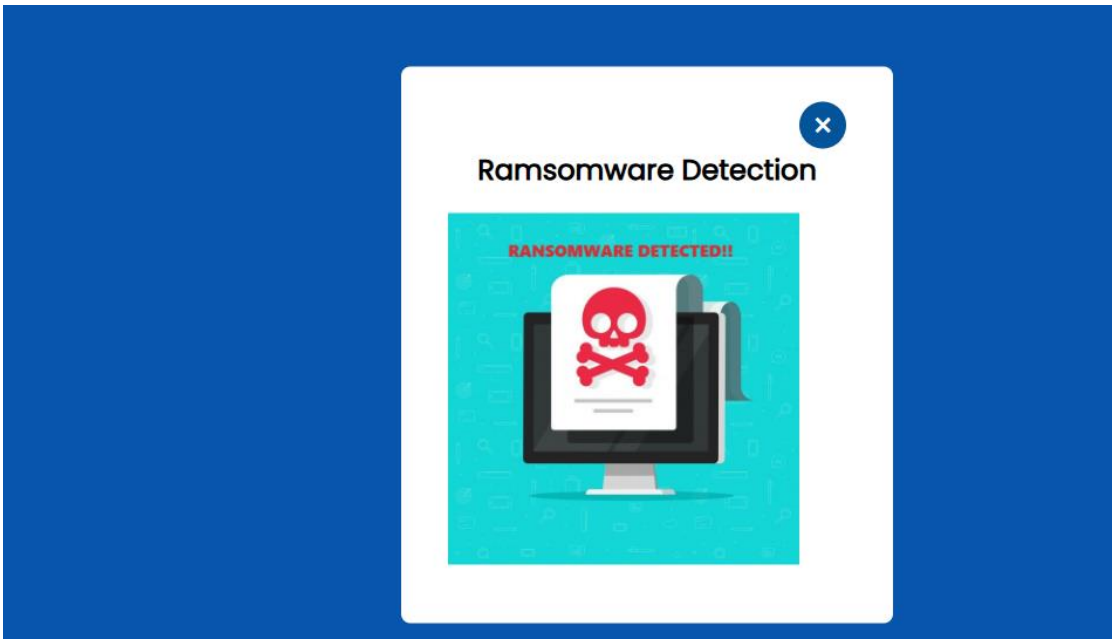

Figure 4: Ransomware Detection

Figure 4 shows the system has successfully detected a ransomware attack, alerting the user to potential data threats.

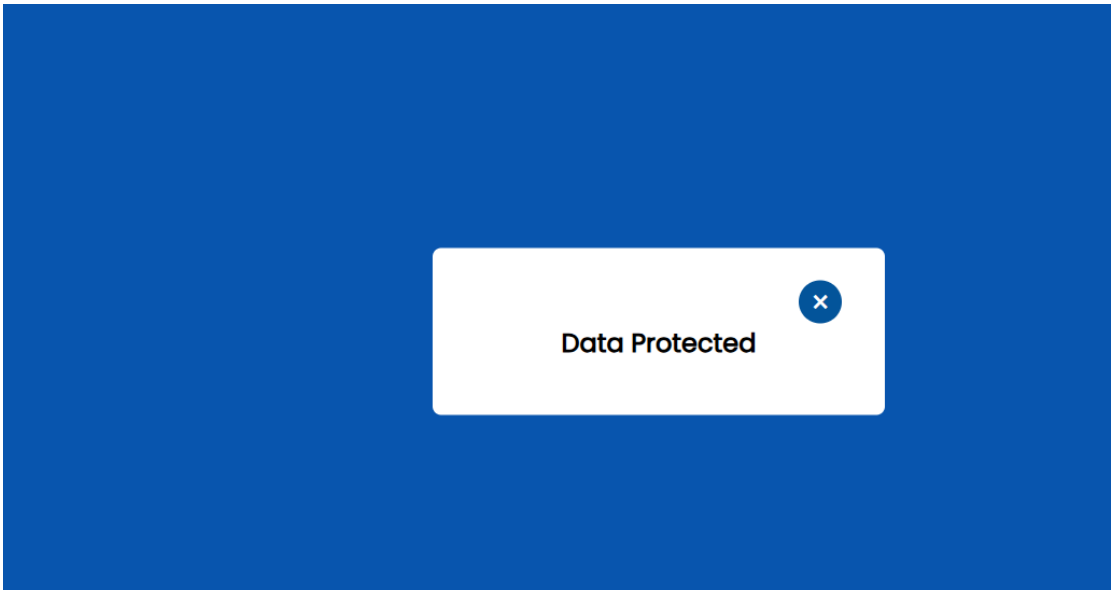

Figure 5: Data Protection

Figure 5 Illustrates once ransomware is detected, the selected files are automatically secured by converting them into a hidden format. This ensures that cybercriminals cannot access or encrypt the protected data.

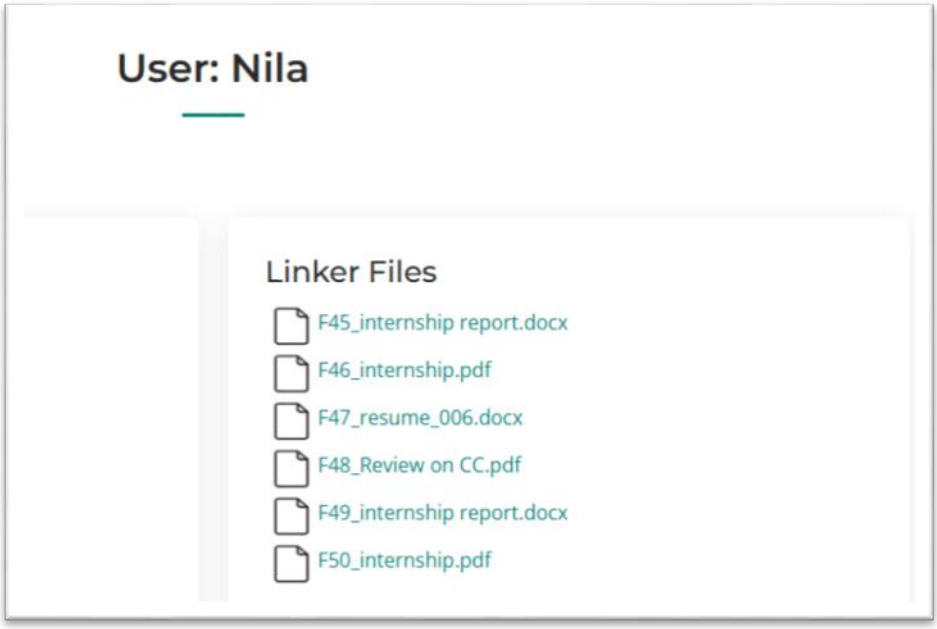

Figure 6: Hidden Layer

Figure 6 displays the hidden layer where the user's selected files are securely stored, contains encoded versions of the original files, making them inaccessible to ransomware.

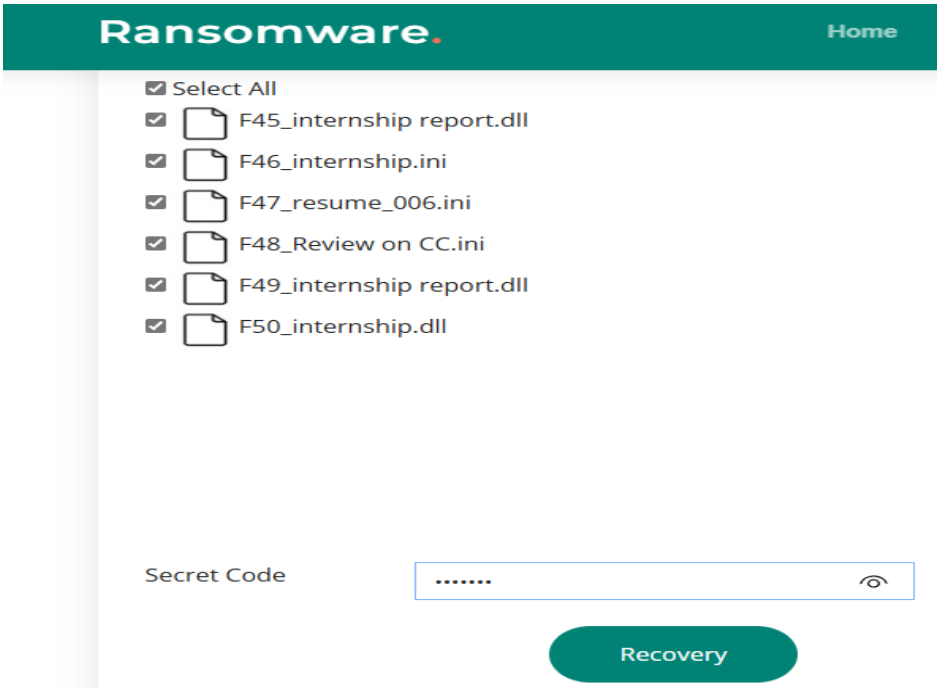

Figure 7: File Recovery

Figure 7 represents the original documents, which can be recovered using a secret code and ensures that even if ransomware infiltrates the system, the protected files remain untouched and retrievable.

4. Conclusion

The proactive defense system against ransomware, integrating BiLSTM and Proactive Data Protection, marks a significant advancement in cybersecurity. By combining BiLSTM's generative capabilities with cryptographic file concealment, the system creates a multi-layered defense that effectively disrupts ransomware operations. Designed for adaptability, it offers real-time monitoring, a user-friendly interface, and customizable security settings, fostering a collaborative approach to ransomware defense. Beyond individual protection, it sets a new industry standard, emphasizing continuous improvement, user education, and cross-platform resilience.

5. References

[1] Smith D, Roy K, 2022. Machine Learning Algorithms and Frameworks in Ransomware Detection. *IEEE Xplore*. [DOI: 10.1109/ACCESS.2022.9934917].

[2] Song Z, Tian Y, Zhang J, 2023. Similarity Analysis of Ransomware Attacks Based on ATT&CK Matrix. *IEEE Xplore*, [DOI: 10.1109/ACCESS.2023.10273710].

[3] Mahesh B, 2020. Machine Learning Algorithms - *A Review.* International Journal of Science and Research (IJSR*), 9(1).* [ISSN: 2319-7064].

[4] Jyothi PNS, Yamagant R, 2019. A Review on Python for Data Science, Machine Learning, and IoT. *International Journal of Scientific & Engineering Research*, 10(12): 851. [ISSN: 2229-5518].

[5] Yadav B, Shrivastava V, Pandey A, Sharma P, 2024. MySQL: Database Design for Performance and Scalability. *International Journal of Research Publication and Reviews*, 5(3): 7631-7638. [ISSN: 2582-7421].