

# **Malware Detection on Devices using Federated Learning in computer devices**

## **DISSERTATION**

**Submitted in partial fulfilment of the requirements of the**

**Degree : MTech in Artificial Intelligence and  
Machine Learning**

**By**

**Sohini Kar**

**2023AB05048**

**Under the supervision of**

**Atin Nandi**

**BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE**

**Pilani (Rajasthan) INDIA**

**August, 2025**

# Acknowledgement

---

I would like to express my sincere gratitude to my mentor for their invaluable guidance, encouragement, and constructive feedback throughout the course of my dissertation. Their insights have been instrumental in shaping the direction and quality of this work.

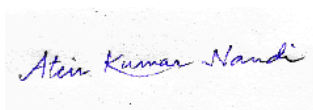
I am deeply thankful to my parents for their constant support, patience, and motivation, which have been a source of strength during this academic journey.

I would also like to acknowledge Tata Consultancy Services (TCS) for providing me the opportunity and support to begin my M.Tech program, and IBM for allowing me to continue and complete my studies while pursuing my professional commitments.

Finally, I extend my appreciation to all those who, directly or indirectly, have contributed to the successful completion of this dissertation

**BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI****CERTIFICATE**

This is to certify that the Dissertation entitled Malware Detection on Devices using Federated Learning in computer devices and submitted by Ms. Sohini Kar ID No. 2023AB05048 in partial fulfilment of the requirements of AIMLCZG628T Dissertation, embodies the work done by him/her under my supervision.



Signature of the Supervisor

Place: Kolkata

Date: 15<sup>th</sup> August 2025

Name: Atin Nandi

Designation: Service Delivery Manager

# Abstract

---

Malware detection systems must achieve high accuracy while safeguarding sensitive user data. Conventional centralized machine learning models often aggregate raw data in a single repository, which increases the risk of information leakage. To address this challenge, this study employs federated learning (FL) for memory-based malware detection and classification. FL enables collaborative model training across distributed clients without requiring direct data sharing, thus preserving privacy. Using the CIC-MalMem-2022 dataset, a total of 22 models were developed and tested, including both feedforward neural networks and long short-term memory architectures, alongside centralized baselines. Experimental results show that FL achieves exceptional performance, with 0.999 accuracy in binary classification and 0.845 in multiclass classification, even under conditions of heterogeneous client data. These outcomes highlight the potential of FL as a secure and privacy-preserving paradigm for malware detection, offering a practical alternative to traditional centralized learning approaches.

## List of Symbols & Abbreviations Used

Index	Abbreviation	Description
1	<b>API</b>	Application Programming Interface
2	<b>BC</b>	Binary Classification
3	<b>BD</b>	Binary Dataset
4	<b>CIC</b>	Canadian Institute for Cybersecurity
5	<b>CIC-MalMem-2022</b>	Memory-based malware dataset from CIC
6	<b>CNN</b>	Convolutional Neural Network
7	<b>DL</b>	Deep Learning
8	<b>DNN</b>	Deep Neural Network
9	<b>FedAvg</b>	Federated Averaging (Aggregation Algorithm)
10	<b>FL</b>	Federated Learning
11	<b>FNN</b>	Feedforward Neural Network
12	<b>GPU</b>	Graphics Processing Unit
13	<b>IID</b>	Independent and Identically Distributed
14	<b>LSTM</b>	Long Short-Term Memory (Recurrent Neural Network)
15	<b>MC</b>	Multiclass Classification
16	<b>MD</b>	Multiclass Dataset
17	<b>ML</b>	Machine Learning
18	<b>NLP</b>	Natural Language Processing
19	<b>RAM</b>	Random Access Memory
20	<b>ReLU</b>	Rectified Linear Unit (activation function)
21	<b>RNN</b>	Recurrent Neural Network
22	<b>SGD</b>	Stochastic Gradient Descent
23	<b>TPU</b>	Tensor Processing Unit

## List of Tables

---

Index	Table No.	Title
<b>1</b>	Table 6.1	Classification Report for Binary Classification (Precision, Recall, F1-Score, Support)
<b>2</b>	Table 6.2	Confusion Matrix for Binary Classification (Model 1)
<b>3</b>	Table 6.3	Confusion Matrix for Binary Classification (Model 2)
<b>4</b>	Table 6.4	Classification Report for Multiclass Classification (FNN-MC)
<b>5</b>	Table 6.5	Classification Report for Multiclass Classification (LSTM-MC)
<b>6</b>	Table 6.6	Experimental Setup and Results of Centralized vs. Federated Learning for FNN and LSTM Models (Binary and Multiclass Classification)

## List of figures

---

Index	Figure No.	Title
1	Figure 3.5	Federated Learning Architecture (Global Model, Local Models, and Aggregation)
2	Figure 6.1	Confusion Matrix for Binary Classification (Model 1)
3	Figure 6.2	Confusion Matrix for Binary Classification (Model 2)
4	Figure 6.3	Confusion Matrix for Multiclass Classification (FNN-MC)
5	Figure 6.4	Confusion Matrix for Multiclass Classification (LSTM-MC)

# Table of Contents

---

<b>Chapter no.</b>	<b>Chapters</b>	<b>Page Numbers</b>
<b>1.</b>	Objectives Met	9
<b>2.</b>	Choosing of Dataset	12
<b>3.</b>	Preprocessing	14
<b>4.</b>	Different Architectures Used	18
<b>5.</b>	Implementation of Federated Learning	24
<b>6.</b>	Results and Discussion	28
<b>7.</b>	Conclusion	37
<b>8.</b>	References	39



# Chapter 1: Objectives met

---

## 1. Literature Review

The primary objective of this section was to survey and critically analyse existing research in the areas of malware detection and federated learning. The review was conducted across two dimensions:

1. **CIC-MalMem-2022-based Malware Detection:**

A number of prior studies have employed the CIC-MalMem-2022 dataset for malware detection; however, these efforts have been predominantly centralized, relying on conventional supervised learning pipelines. The review revealed that while promising accuracies were achieved, these approaches did not address the inherent privacy challenges associated with handling sensitive system data, nor did they explore distributed frameworks for secure model training.

2. **Federated Learning (FL)-based Malware Detection:**

A parallel strand of literature was identified where federated learning had been applied to cybersecurity tasks, but on datasets other than CIC-MalMem-2022. These studies demonstrated the potential of FL in privacy-preserving learning, particularly in malware and intrusion detection scenarios.

The literature review thus established a **research gap**: the lack of exploration of federated learning using the CIC-MalMem-2022 dataset. This highlighted the novelty and significance of this work, motivating the development of a federated framework specifically tailored for memory-dump-based malware detection.

---

## 2. Dataset Selection and Preprocessing

This section addressed the critical objective of preparing the dataset for effective experimentation. The **CIC-MalMem-2022 dataset** was selected due to its comprehensive coverage of benign and malicious Windows processes captured through memory dumps, providing realistic and fine-grained system-level behavioral data.

The preprocessing workflow included:

- **Binary Classification Setup:** The *Category* column was converted into binary labels {0, 1}, mapping benign vs. malicious processes, to align with `binary_crossentropy` as the loss function.
- **Multiclass Classification Setup:** Malware families such as ransomware, spyware, and trojan horse were encoded into numerical indices {0, 1, 2, 3}, enabling `sparse_categorical_crossentropy` for multiclass tasks.
- **Feature Normalization:** All numerical features were scaled to a normalized range, ensuring stable gradient descent optimization across architectures.
- **LSTM-specific Reshaping:** Since LSTMs require sequential input, the feature vectors were reshaped to a three-dimensional format (*samples, 1, features*).

- **String Cleaning and Label Consistency:** Malware identifiers with long hash-based suffixes were truncated to retain only lowercase family names (e.g., “ransomware”, “trojan horse”, “spyware”), ensuring semantic clarity and label consistency.

Through this preprocessing, the dataset was systematically structured to support both centralized and federated learning experiments across binary and multiclass classification tasks.

---

### 3. Model Architecture Design

The objective of this section was to design robust deep learning architectures capable of learning from memory-dump features under both centralized and federated training conditions. Four architectures were constructed:

- **FNN-BC:** A feedforward neural network tailored for binary classification tasks.
- **FNN-MC:** A feedforward neural network adapted for multiclass classification tasks.
- **LSTM-BC:** A sequential Long Short-Term Memory network for binary classification, leveraging temporal dependencies in features.
- **LSTM-MC:** An LSTM network extended for multiclass classification.

Each architecture was developed as a modular build function, allowing flexible reuse in centralized and federated pipelines. The models incorporated **Rectified Linear Unit (ReLU)** activation functions, dropout regularization (where applicable), and adaptive optimizers to maximize performance.

This systematic design ensured architectural comparability across experiments while capturing both static (via FNNs) and sequential (via LSTMs) feature relationships.

---

### 4. Federated Learning Setup

The objective of this section was to establish a fully functional federated learning pipeline to compare against centralized baselines. A custom **make\_federate** function was implemented, encapsulating all aspects of federated model training.

Key components included:

- **Input Configuration:** Accepted dataset splits, labels, and architecture choice.
- **Training Parameters:** Number of users (clients), local epochs, batch sizes, and stopping criteria based on target accuracy.
- **Optimizer:** Adam optimizer with hyperparameters (learning rate = 0.001,  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,  $\epsilon = 1e-08$ ).
- **Data Partitioning:** IID splits were generated using a custom **data\_shuffle** function, ensuring fairness across clients.
- **Federated Aggregation:** Local model weights were trained independently and combined via **Federated Averaging (FedAvg)**, where updates were scaled by client data size before averaging.

- **Evaluation and Early Stopping:** After each communication round, the aggregated global model was evaluated, with training terminating once the target accuracy was achieved.

Controlled experiments were performed with varying client counts (8 and 16), enabling observation of **performance trade-offs** in terms of accuracy, convergence time, and training efficiency. This validated the robustness of the proposed federated system.

---

## 5. Optimization and Evaluation

The final objective was to rigorously optimize and evaluate the models under both centralized and federated settings. The strategies applied included:

- **Loss Functions:** Binary cross-entropy for binary tasks, and sparse categorical cross-entropy for multiclass tasks.
- **Mini-Batch Gradient Descent:** Adopted for improved convergence stability, reduced gradient variance, and efficient GPU utilization.
- **Performance Metrics:** Accuracy, precision, recall, and F1-score were employed to provide a balanced evaluation of models. These metrics ensured both per-class and overall performance insights.
- **Result Logging and Visualization:** Experimental outcomes were systematically recorded and plotted, enabling comparative analysis between centralized vs. federated approaches and between different architectures.

The evaluation confirmed that while federated models achieved comparable performance to centralized ones, they also introduced **notable privacy-preserving advantages** with respect to data sharing. This validated the overarching objective of deploying federated learning for malware detection using memory-dump features.

## Chapter 2: Related Work and Dataset Selection

---

### 2.1 Introduction

The detection of malware has long been a critical challenge in the field of cybersecurity. Traditional detection mechanisms, primarily based on static signatures and heuristic rules, are increasingly inadequate in the face of rapidly evolving threats such as ransomware, spyware, and trojan families. Machine learning (ML) and deep learning (DL) techniques have therefore emerged as powerful alternatives, capable of identifying complex behavioural patterns within system data. However, most conventional ML approaches rely on **centralized training** frameworks, which require the collection of sensitive data in a single location. This introduces significant risks of **data leakage, privacy violations, and regulatory non-compliance**.

Federated Learning (FL) offers a promising solution to this challenge by allowing distributed clients to collaboratively train models without sharing raw data. Instead, only model updates (e.g., weights and gradients) are exchanged, thereby ensuring privacy and compliance with data protection principles. This chapter provides the background required to situate the present research in the broader field. It first surveys related work in malware detection using both centralized and federated paradigms. It then presents the details of dataset preprocessing for the CIC-MalMem-2022 dataset, which forms the foundation of the experimental setup in later chapters.

---

### 2.2 Related Work

#### 2.2.1 Centralized Approaches to Malware Detection

Research on malware detection has historically relied on centralized approaches, wherein all data samples are aggregated into a central repository for training. Studies leveraging the **CIC-MalMem-2022 dataset** have applied classical algorithms such as **Random Forests, Support Vector Machines (SVMs), and Decision Trees**, as well as deep learning models including **Feedforward Neural Networks (FNNs)** and **Convolutional Neural Networks (CNNs)**.

These methods demonstrated high predictive accuracy, with some works achieving over 95% detection rates for binary classification tasks. However, their dependence on centralized data posed challenges:

- **Scalability:** As the volume of system-level data grows, centralized storage becomes computationally expensive.
- **Privacy Risks:** Malware detection datasets often originate from enterprise or personal computing environments, where sensitive memory features could expose confidential information.
- **Vulnerability to Breaches:** Data aggregation increases the attack surface, making centralized repositories attractive targets for adversaries.

Thus, while centralized approaches provided strong baselines, their practical deployment in privacy-sensitive environments remained limited.

---

### 2.2.2 Federated Learning in Cybersecurity

The advent of FL has opened new avenues in the design of **privacy-preserving malware detection systems**. Early studies explored FL in the context of **Android malware detection**, where device-level data such as permissions, API calls, and application behaviors were distributed across multiple clients. Similarly, FL frameworks have been applied to **IoT environments**, leveraging distributed network traffic to collaboratively train anomaly detection models without exposing raw data.

These studies highlighted several strengths of FL:

1. **Privacy Preservation:** Sensitive features never leave the client device, thereby minimizing the risk of leakage.
2. **Collaborative Intelligence:** Multiple clients contribute to building a more generalized global model.
3. **Regulatory Compliance:** FL aligns with privacy regulations such as GDPR and HIPAA, which restrict raw data sharing.

Despite these advantages, the application of FL to **system-level memory dump datasets**—such as CIC-MalMem-2022—remains underexplored. Existing research has mostly focused on mobile or network traffic datasets, which differ in feature representation and attack visibility. System memory, however, captures fine-grained process behavior, offering a unique lens for detecting sophisticated malware.

---

### 2.2.3 Identified Research Gap

From the above survey, two important gaps emerge:

- The **CIC-MalMem-2022 dataset** has only been analyzed in **centralized settings**, limiting its potential for deployment in real-world distributed systems.
- **Federated Learning frameworks** have been widely applied to alternative datasets (e.g., Android apps, IoT traffic), but their effectiveness on **memory dump features** remains unexplored.

This research directly addresses these gaps by applying **Federated Learning to CIC-MalMem-2022**, thereby combining the strengths of privacy-preserving training with a dataset that captures realistic system-level malware behavior.

---

## 2.3 Dataset Preprocessing

### 2.3.1 Overview of CIC-MalMem-2022

The CIC-MalMem-2022 dataset, developed by the Canadian Institute for Cybersecurity, provides a comprehensive benchmark for evaluating malware detection techniques. It consists of **memory dump features** extracted from both **benign and malicious Windows processes**, encompassing diverse families such as ransomware, spyware, and trojan horse malware.

The dataset includes **57 attributes**, reflecting process-level memory behaviour. However, not all attributes were informative for model training, necessitating rigorous preprocessing before experimentation. The preprocessing steps aimed to generate two refined datasets:

- **Binary Dataset (BD):** Designed for distinguishing benign from malicious processes.
- **Multiclass Dataset (MD):** Designed for classifying processes into benign, ransomware, spyware, or trojan categories.

### 2.3.2 Data Cleaning

The preprocessing began with the removal of redundant attributes. Columns with **constant values across all rows** (IDs 5, 11, and 52) were eliminated, reducing dimensionality to 54 features. Additionally, the **Category column** contained long identifiers combining malware family names with unique hash values (e.g., “*Ransomware-Shade-fa03be...*”). These were simplified to retain only the **family name**, improving readability and consistency.

Basic data quality checks were performed to ensure there were **no missing values, infinite values, or corrupted entries** in the dataset. This ensured reliability in downstream training tasks.

### 2.3.3 Binary Dataset Preparation

The **Binary Dataset (BD)** was constructed by:

1. Retaining the **Class attribute** as the target variable (*benign*  $\rightarrow$  0, *malware*  $\rightarrow$  1).
2. Removing the **Category attribute**, since its values could trivially determine the class label, leading to information leakage.
3. Standardizing the 52 independent variables using **Z-score normalization**, ensuring uniform feature scaling.

This process yielded a dataset with **53 variables** (52 features + 1 target). The BD was thus suitable for binary classification tasks using both centralized and federated models.

### 2.3.4 Multiclass Dataset Preparation

The **Multiclass Dataset (MD)** was created from a copy of the slightly modified dataset by:

1. Using the **Category attribute** as the target variable, with mappings:
  - Benign  $\rightarrow$  0
  - Ransomware  $\rightarrow$  1
  - Spyware  $\rightarrow$  2
  - Trojan Horse  $\rightarrow$  3
2. Removing the **Class attribute** to avoid redundancy.
3. Applying **Z-score normalization** to all independent variables.

This produced a dataset with 52 independent features and one categorical target variable.

---

### 2.3.5 Data Reshaping for Neural Networks

To prepare the datasets for deep learning:

- **Feedforward Neural Networks (FNNs):** Data was retained in its two-dimensional form (*samples, features*).
- **Long Short-Term Memory (LSTM) Networks:** Data was reshaped into three dimensions (*samples, 1, features*) to satisfy sequential input requirements.

This ensured compatibility across different model architectures.

---

### 2.3.6 Preprocessing Workflow

The entire preprocessing pipeline can be summarized as follows:

1. Load CIC-MalMem-2022 dataset (57 features).
  2. Remove redundant attributes ( $\downarrow$  to 54 features).
  3. Clean Category column (retain family names).
  4. Create **Binary Dataset (BD)**: Keep *Class*, drop *Category*.
  5. Create **Multiclass Dataset (MD)**: Keep *Category*, drop *Class*.
  6. Apply Z-score normalization.
  7. Reshape for FNN and LSTM input formats.
-

## 2.4 Summary

This chapter presented the **related work** in malware detection and federated learning, highlighting the lack of studies that apply FL to the CIC-MalMem-2022 dataset. It also detailed the **dataset preprocessing steps**, including cleaning, transformation, normalization, and preparation of binary and multiclass variants.

Through these steps, the dataset was made compatible with both **traditional feedforward architectures** and **sequence-based LSTM models**, forming the basis for the experimental design presented in the next chapter.



## Chapter 3: Preprocessing of Dataset

---

### 3.1 Introduction

Data preprocessing is a critical step in any machine learning pipeline, directly influencing model performance and generalization capability. The CIC-MalMem-2022 dataset, although well-structured and comprehensive, required several preprocessing operations before being utilized for training binary and multiclass malware classification models under the proposed federated learning framework. These preprocessing steps served two primary purposes:

1. To prepare two separate datasets — one for **binary classification** and one for **multiclass classification** — derived from a slightly modified version of the original dataset.
2. To ensure that the input features were consistent, relevant, and appropriately scaled for deep learning architectures, thereby improving training stability and convergence.

---

### 3.2 Initial Data Cleaning and Attribute Reduction

The original CIC-MalMem-2022 dataset consists of 57 attributes, including system and memory usage statistics captured during both benign and malicious activities. The first preprocessing step involved **removing attributes with no discriminative value** — specifically, columns where all entries contained the same value across all rows. Such columns provide no useful information to the learning algorithm and may add unnecessary noise.

From the dataset, column names with `pslist.nprocs64bit`, `handles.nport` and `svcsan.interactive_process_services`, were identified as having constant values and were therefore removed. This reduction decreased the total number of attributes from 57 to **54**.

Following attribute reduction, basic data integrity checks were performed to:

- Detect and handle **infinite values** in any column.
- Verify that there were **no missing values** across rows.

The results of these checks indicated that the dataset was free of NaN (Not a Number) entries and infinite values, allowing preprocessing to continue without imputation.

---

### 3.3 Target Variable Refinement and Label Encoding

#### 3.3.1 Modifying the Category Column

The *Category* column in the original dataset contained descriptive strings that combined malware family names with unique identifiers, for example:

```
Ransomware-Shade-fa03be3078d1b9840f06745f160eb...
```

Such identifiers were removed so that the column retained only the **malware family name** in lowercase (e.g., *ransomware*, *trojan horse*, *spyware*). This transformation ensured consistent labelling and eliminated extraneous string elements that could not contribute to learning.

#### 3.3.2 Mapping the Class Column for Binary Classification

The *Class* column distinguished between *benign* and *malware* entries. For binary classification, this column was encoded numerically:

- **benign** → 0
- **malware** → 1

This binary encoding was essential for compatibility with the **binary\_crossentropy** loss function used in binary classification models.

---

### 3.4 Dataset Preparation for Binary Classification (BD)

A **Binary Dataset (BD)** was created from the slightly modified original dataset. The key steps were:

1. **Removal of the Category Column** – The *Category* attribute, which contains highly predictive information about the target variable, was removed to prevent data leakage. Retaining this attribute would have allowed the model to trivially classify entries without learning from the other 52 features, undermining the purpose of the study.
2. **Target Variable** – The *Class* column was retained as the prediction target.
3. **Standardization** – All 52 independent variables (excluding the target) were standardized using **Z-score normalization**, ensuring that each feature had zero mean and unit variance:

$$z = (x - \mu) / \sigma$$

This normalization is known to improve convergence speed and stability in neural networks by preventing features with larger numerical ranges from dominating the learning process.

After preprocessing, BD contained **53 variables** (52 features + 1 target) and preserved a balanced distribution between the two target classes — **29,298 benign** and **29,298 malware** samples.

### 3.5 Dataset Preparation for Multiclass Classification (MD)

A **Multiclass Dataset (MD)** was prepared using a similar procedure but with modifications appropriate for multiclass prediction:

1. **Removal of the Class Column** – In multiclass classification, the target variable is *Category*, which contains four classes:
  - Benign (0)
  - Ransomware (1)
  - Spyware (2)
  - Trojan Horse (3)

The *Class* column was removed to avoid redundancy and potential data leakage.

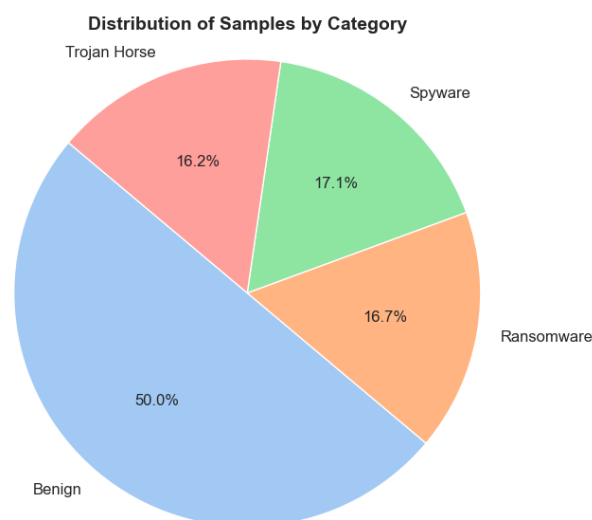
2. **Target Variable** – The *Category* column values were label-encoded into integers {0, 1, 2, 3}.
3. **Standardization** – As with BD, Z-score normalization was applied to all 52 independent variables.

The final MD dataset contained 52 features and 1 categorical target variable, with the class distribution shown below:

**Table 3.5.1:** Class Distribution and Label Encoding of the CIC-MalMem-2022 Dataset

Category	Samples	Encoded Label
Benign	29,298	0
Ransomware	9,791	1
Spyware	10,020	2
Trojan Horse	9,487	3

**Figure 3.5.1:** Distribution of CIC-MalMem-2022 Samples by Category



---

### 3.6 Dataset Compatibility with Model Architectures

Both BD and MD were prepared for direct integration with the study's deep learning models:

- **Binary Classification Models (FNN-BC, LSTM-BC)** – Used BD with `binary_crossentropy` loss.
- **Multiclass Classification Models (FNN-MC, LSTM-MC)** – Used MD with `sparse_categorical_crossentropy` loss.

For LSTM architectures, additional preprocessing was applied:

- The 2D feature matrix was reshaped into a 3D tensor of shape **(samples, 1, features)**, treating each sample as a sequence of length 1 with 52 features. This structure is required for Keras/TensorFlow LSTM layers.

---

### 3.7 Summary of Preprocessing Workflow

The preprocessing pipeline can be summarized in the following steps:

1. Remove constant-value columns (IDs 5, 11, 52).
2. Verify dataset integrity (no missing or infinite values).
3. Clean *Category* column by removing unique identifiers.
4. Encode *Class* for binary classification (0 = benign, 1 = malware).
5. Create BD by removing *Category*, keeping *Class* as target.
6. Create MD by removing *Class*, encoding *Category* as target.
7. Apply Z-score normalization to all feature columns.
8. Reshape data for LSTM-based models where applicable.

## Chapter 4: Model Architectures for Binary and Multiclass Classification

### 4.1 Introduction

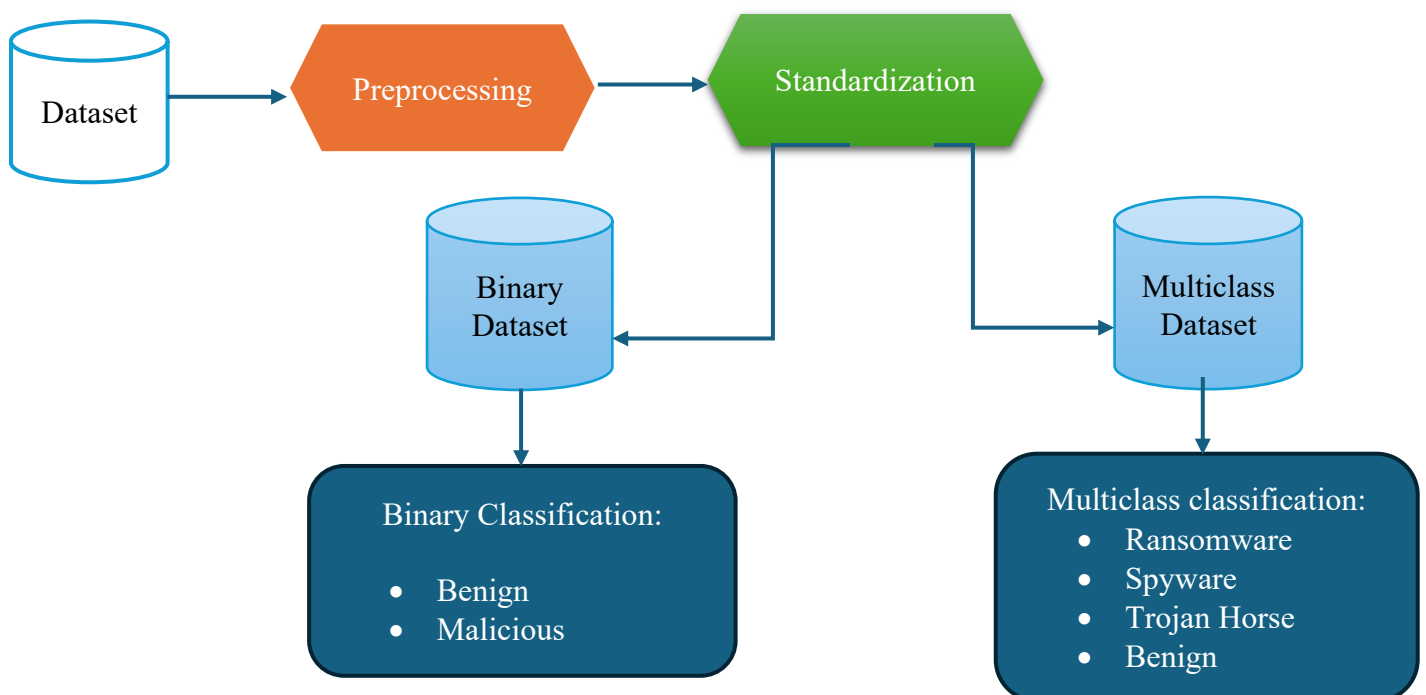
This chapter presents the neural network architectures designed and implemented in this study for malware detection, addressing both **binary classification** (benign vs. malicious) and **multiclass classification** (identifying specific malware families). Two primary deep learning paradigms were employed — **Feedforward Neural Networks (FNN)** and **Long Short-Term Memory Networks (LSTM)** — each adapted to the specific requirements of binary and multiclass problems.

The four proposed base architectures are:

- **FNN-BC:** Feedforward Neural Network for Binary Classification
- **FNN-MC:** Feedforward Neural Network for Multiclass Classification
- **LSTM-BC:** Long Short-Term Memory Network for Binary Classification
- **LSTM-MC:** Long Short-Term Memory Network for Multiclass Classification

Each architecture was implemented in **centralized (non-federated)** and **federated learning (FL)** modes, producing a total of 22 base models. Additional experimental variants were created to explore the influence of architectural and training parameter changes. This chapter describes the underlying algorithms, design rationale, architectural specifications, and suitability of each model for deployment in a federated environment.

Figure 4.1.1: Dataset Preprocessing and Classification Workflow



---

## 4.2 Deep Learning Algorithms

### 4.2.1 Feedforward Neural Networks (FNN)

Feedforward Neural Networks are the most widely adopted class of neural architectures for supervised learning. They are composed of an input layer, one or more hidden layers, and an output layer, with information flowing strictly in the forward direction. Each layer performs a linear transformation followed by a non-linear activation function, enabling the network to approximate complex, non-linear relationships in the data.

FNNs are computationally efficient, making them suitable for real-time and resource-constrained environments. In this study, FNNs form the basis for both binary and multiclass classification tasks, with architectural adjustments to layer sizes, activation functions, and dropout rates to optimize performance.

---

### 4.2.2 Long Short-Term Memory Networks (LSTM)

Long Short-Term Memory networks extend the capabilities of Recurrent Neural Networks (RNNs) by addressing the vanishing gradient problem through the introduction of **gated memory units**. These gates — input, forget, and output — regulate the retention and updating of information, allowing the network to capture long-range dependencies.

Although traditionally used for sequential data such as text or time series, LSTMs are also effective in modeling dependencies within non-sequential structured data. For the CIC-MalMem-2022 dataset, LSTMs provide the ability to learn complex inter-feature dependencies that may not be captured by a purely feedforward approach.

---

## 4.3 Design Considerations

The design of each architecture considered the following factors:

### 1. Activation Functions

- **Binary classification (FNN-BC, LSTM-BC):** *Rectified Linear Unit (ReLU)* in hidden layers for computational efficiency and improved gradient propagation, with *Sigmoid* in the output layer for probabilistic binary prediction.
- **Multiclass classification (FNN-MC, LSTM-MC):** *Tanh* in hidden layers to normalize activations between -1 and 1, with *Softmax* in the output layer to generate class probability distributions.

### 2. Regularization via Dropout

Dropout layers deactivate a fraction of neurons during training to prevent overfitting. Rates were tuned per architecture:

- 0.5 for binary models (FNN-BC, LSTM-MC where applicable)
- 0.4 for multiclass FNN models (FNN-MC)

### 3. Model Complexity and FL Constraints

Since federated learning involves client devices that may have limited computational

resources, architectures were designed to be lightweight while maintaining high classification performance. This balance was achieved by minimizing the number of layers and neurons without compromising accuracy.

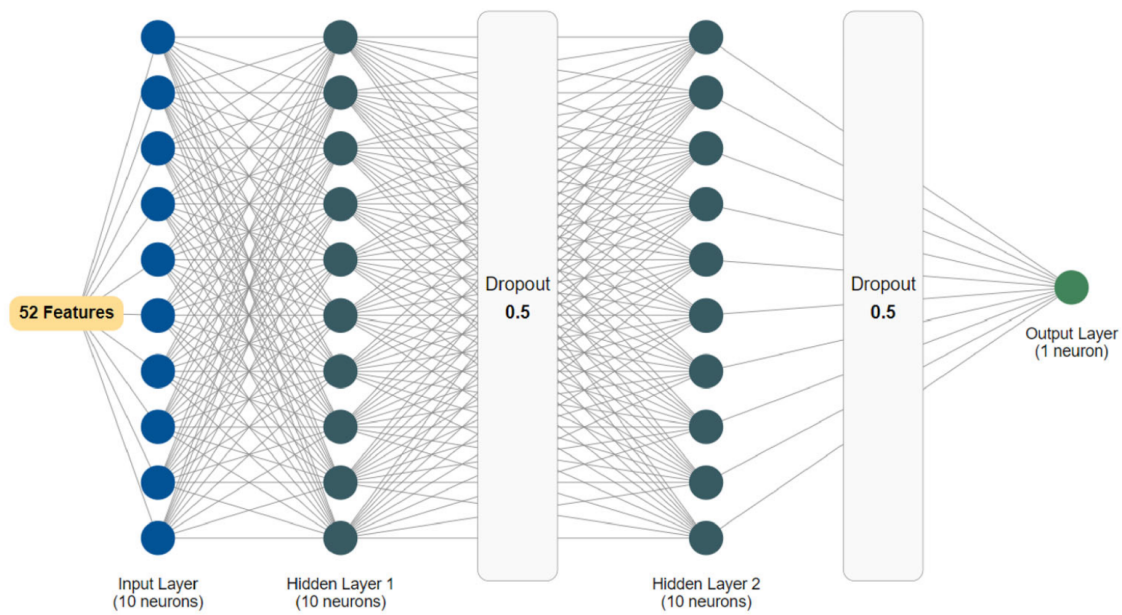
## 4.4 Proposed Architectures

### 4.4.1 FNN-BC: Feedforward Neural Network for Binary Classification

- **Input Layer:** 52 features  $\rightarrow$  Dense(10) with ReLU activation
- **Hidden Layer 1:** Dense(10, ReLU) + Dropout(0.5)
- **Hidden Layer 2:** Dense(10, ReLU) + Dropout(0.5)
- **Output Layer:** Dense(1, Sigmoid)

This compact architecture is optimized for speed and low resource usage in federated settings, while dropout mitigates overfitting.

**Figure 4.4.1.1: Architecture of Feedforward Neural Network for Binary Classification (FNN-BC)**



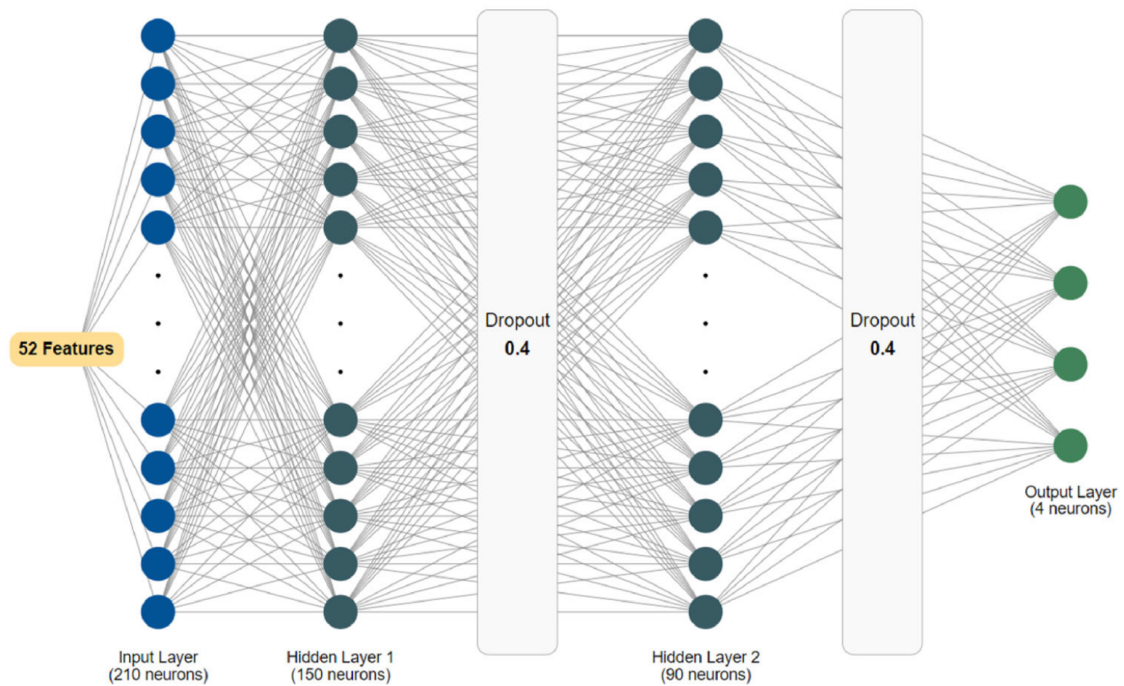
### 4.4.2 FNN-MC: Feedforward Neural Network for Multiclass Classification

- **Input Layer:** 52 features  $\rightarrow$  Dense(210, Tanh)
- **Hidden Layer 1:** Dense(150, Tanh) + Dropout(0.4)
- **Hidden Layer 2:** Dense(90, Tanh) + Dropout(0.4)
- **Output Layer:** Dense(4, Softmax)

The larger neuron counts and dropout rates reflect the increased complexity of multiclass classification.



**Figure 4.4.2.1: Architecture of Feedforward Neural Network for Multiclass Classification (FNN-MC)**

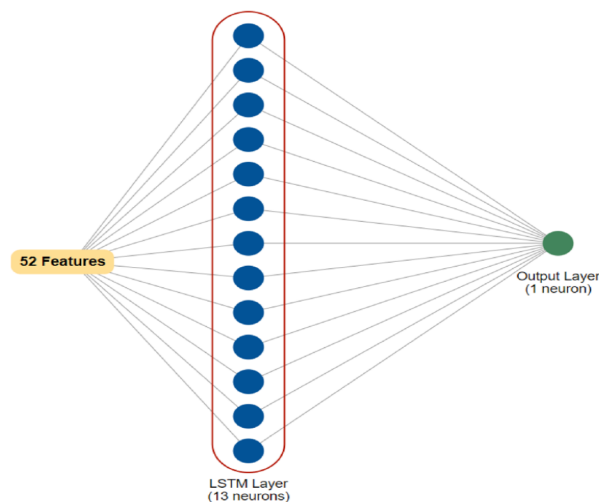


#### 4.4.3 LSTM-BC: Long Short-Term Memory for Binary Classification

- **Input Layer:** LSTM(13 units, ReLU activation)
- **Output Layer:** Dense(1, Sigmoid)

A minimalist LSTM architecture without dropout, chosen to minimize computation time on low-power devices.

**Figure 4.4.3.1: Architecture of Long Short-Term Memory Network for Binary Classification (LSTM-BC)**



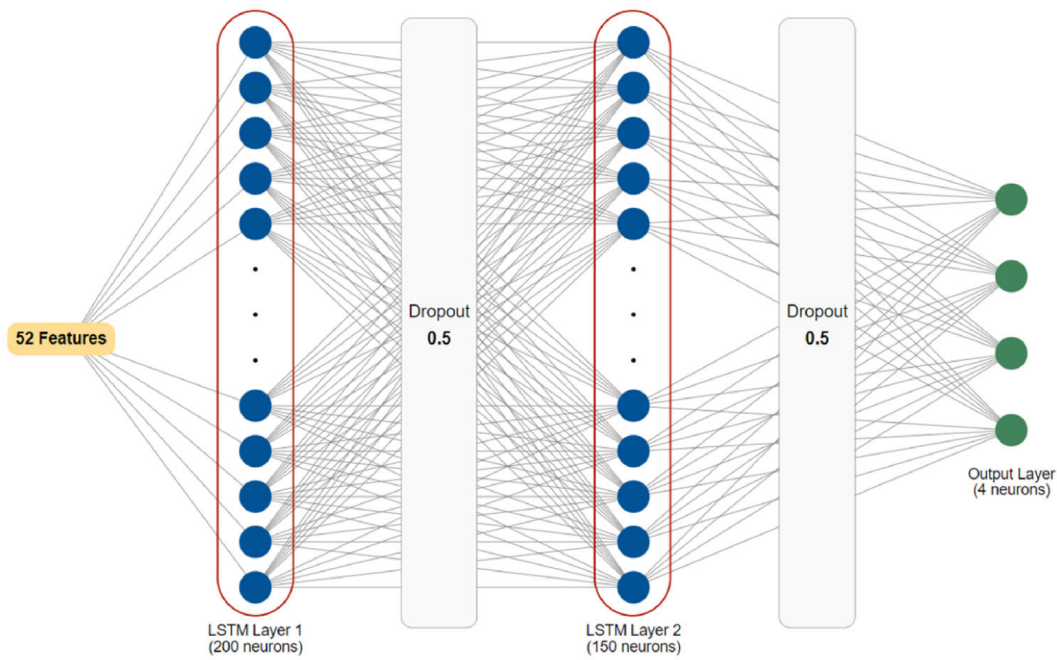


#### 4.4.4 LSTM-MC: Long Short-Term Memory for Multiclass Classification

- **LSTM Layer 1:** LSTM(200 units, Tanh) + Dropout(0.5)
- **LSTM Layer 2:** LSTM(150 units, Tanh) + Dropout(0.5)
- **Output Layer:** Dense(4, Softmax)

This deeper LSTM configuration captures richer inter-feature dependencies necessary for multiclass discrimination.

**Figure 4.4.4.1: Architecture of Long Short-Term Memory Network for Multiclass Classification (LSTM-MC)**



#### 4.5 Federated and Centralized Model Variants

Each architecture was implemented in:

- **Centralized mode:** All training data processed on a single server (users = 1).
- **Federated mode:** Data partitioned across simulated clients with 8, 16, 32, 64, and (for binary models) 128 clients. For multiclass FL experiments, the 128-user scenario was excluded due to excessive training time.

## 4.6 Summary

The four base architectures — FNN-BC, FNN-MC, LSTM-BC, and LSTM-MC — balance performance with computational efficiency, making them suitable for deployment in both centralized and federated environments. Their design reflects trade-offs between accuracy, generalization, and training time, with architecture complexity increasing from binary to multiclass tasks and from FNN to LSTM.

## Chapter 5: Implementation of Federated learning

### 5.1 Introduction

This chapter describes the implementation of the proposed **Federated Learning (FL)**-based malware detection framework, **FEDetect**, developed to enable local clients to perform both malware detection and malware type classification on their private data without compromising confidentiality. The implementation leverages a controlled simulation environment in which both client devices and the FL server are virtualized, allowing for systematic evaluation of the FL process under repeatable conditions.

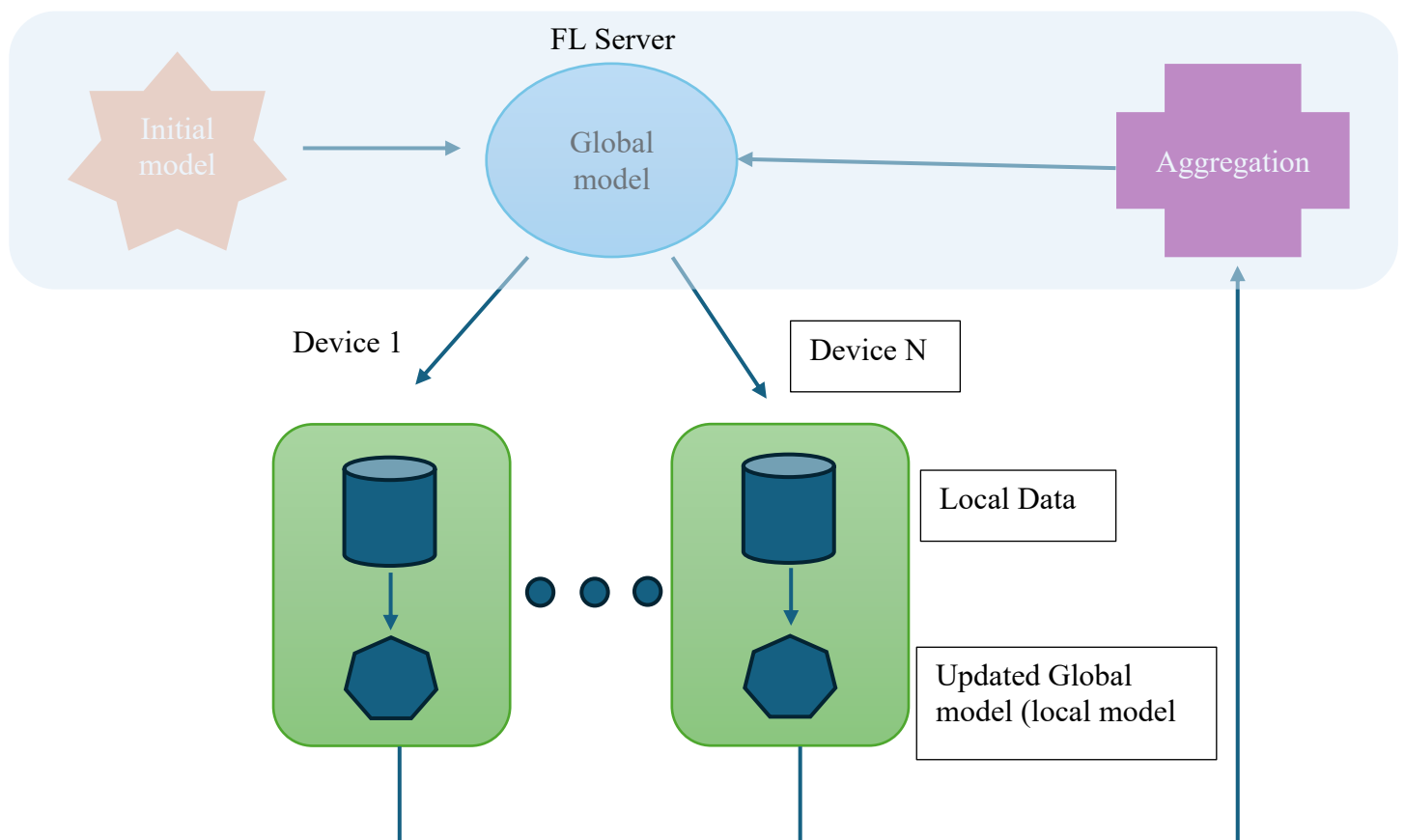
### 5.2 Simulation Environment

#### 5.2.1 Concept and Design

FEDetect is designed to represent a realistic FL ecosystem in which each participating local device trains a model on its private dataset and communicates only model parameters to a central FL server. No raw data leaves the client devices, significantly reducing the risk of privacy breaches. The FL server aggregates model updates from all clients to produce a new global model, which is subsequently redistributed to clients for the next training round.

The simulation environment mimics this workflow entirely on a single high-performance workstation, representing multiple clients and the central server. There is no direct device-to-device communication; all coordination is mediated by the simulated server.

**Figure 5.2.1.1: Federated Learning Architecture**



---

### 5.2.2 Execution Workflow

The FEDetect execution process follows five main steps:

1. **Initialization:**  
The FL server creates an initial model, termed the *Global Model*.
2. **Distribution:**  
The global model is sent to each virtual client sequentially.
3. **Local Training:**  
Each client trains the received model on its local dataset for a specified number of epochs.
4. **Parameter Return:**  
Clients send the updated model weights (local models) back to the FL server.
5. **Aggregation:**  
The server aggregates local models using the **Federated Averaging (FedAvg)** algorithm, producing a new global model.

One iteration of these five steps constitutes a **round**. The process repeats until the target accuracy is achieved or a predefined stopping criterion is met.

---

### 5.3 Model Aggregation Algorithm

The **FedAvg** algorithm was adopted for model aggregation. In this approach, each client's local model weights are averaged to generate the updated global model weights, proportionally balancing contributions across clients.

Unlike traditional implementations that primarily use the **Stochastic Gradient Descent (SGD)** optimizer, this study integrates **Adam** with FedAvg (**FedAvg-Adam**), motivated by recent literature reporting improved convergence and accuracy for Adam in federated contexts.

---

### 5.4 Hardware and Software Setup

- **Hardware:**
  - CPU: Intel Core i7-12650H @ 2.30 GHz
  - RAM: 16 GB
  - Storage: 512 GB SSD
  - GPU: NVIDIA GeForce RTX 3060
- **Software & Libraries:**
  - OS: Windows 10
  - Development Environment: Jupyter Notebook v6.5.2
  - Programming Language: Python v3.11.3
  - Libraries: TensorFlow, Keras, NumPy, Pandas, Scikit-learn

---

## 5.5 Proposed Federated Learning Method

### 5.5.1 The `make_federate` Function

All FL experiments are executed through the `make_federate` function, which encapsulates dataset splitting, client simulation, local training, weight aggregation, and model evaluation.

**Parameters include:**

- **X, y:** Input features and target labels
  - **arc:** Model architecture (`FNN_BC`, `FNN_MC`, `LSTM_BC`, `LSTM_MC`)
  - **model\_num:** Identifier for the model variant
  - **users:** Number of simulated FL clients
  - **acc:** Target accuracy for termination
  - **local\_epochs:** Number of epochs per client per round
  - **optimizer:** Optimizer instance (Adam)
  - **ts:** Test data fraction (0.2 in all experiments)
  - **bs:** Batch size
- 

### 5.5.2 Optimizer and Loss Functions

- **Optimizer:** Adam (`learning_rate` = 0.001, `beta_1` = 0.9, `beta_2` = 0.999, `epsilon` =  $1e-08$ )
- **Loss Functions:**
  - Binary models (FNN-BC, LSTM-BC): `binary_crossentropy`
  - Multiclass models (FNN-MC, LSTM-MC): `sparse_categorical_crossentropy`

Loss function selection aligns with preprocessing outputs: binary targets (0, 1) for binary classification, and integer-encoded class indices (0–3) for multiclass.

---

### 5.5.3 Data Partitioning Strategy

Two data distribution strategies were tested:

- **IID:** Preserves class proportions across all clients.
- **Non-IID:** Produces imbalanced distributions, where some clients may lack certain classes entirely.

Unless otherwise specified, the IID setting was used to represent optimal operating conditions.

---

#### 5.5.4 Training and Evaluation

- After each round, the updated global model is evaluated on the reserved test set.
  - Evaluation metrics: Accuracy, Precision, Recall, F1-score.
  - Results are logged per round into history files for later analysis and visualization.
  - For LSTM-based models, test data are reshaped into `(samples, time_steps=1, features)` format before evaluation.
- 

#### 5.5.5 Mini-Batch Training

Batch size (`bs`) controls how data is split into mini-batches during training, balancing computational efficiency, gradient stability, and generalization capability. Mini-batch training is especially beneficial for large models or datasets under limited memory.

---

### 5.6 Summary

The implemented FL framework, FEDetect, successfully simulates privacy-preserving malware detection using both binary and multiclass neural network architectures. The `make_federate` function provides a flexible, parameter-driven way to experiment with different architectures, data distributions, and client configurations. By combining **FedAvg-Adam** aggregation, carefully chosen loss functions, and both IID and Non-IID data splits, FEDetect creates a robust experimental foundation for evaluating federated malware detection performance.

## Chapter 6: Results and Discussion

This chapter presents the experimental findings of the proposed **FEDetect** system and interprets them in the context of the objectives defined in Chapter 1. The results are compared with traditional centralized models to assess the feasibility and performance trade-offs of using Federated Learning (FL) for malware detection.

### 6.1 Experimental Setup Recap

The experiments used the **CIC-MalMem-2022** dataset, prepared in two formats:

- **Binary Dataset (BD)** for malware vs. benign classification.
- **Multiclass Dataset (MD)** for ransomware, spyware, trojan horse, and benign classification.

Four neural network architectures were evaluated:

- **FNN-BC** – Feedforward Neural Network for Binary Classification.
- **FNN-MC** – Feedforward Neural Network for Multiclass Classification.
- **LSTM-BC** – Long Short-Term Memory Network for Binary Classification.
- **LSTM-MC** – Long Short-Term Memory Network for Multiclass Classification.

Each model was tested in **centralized** and **federated** settings using the **FedAvg-Adam** aggregation approach.

### 6.2 Performance of Centralized Models (Baseline)

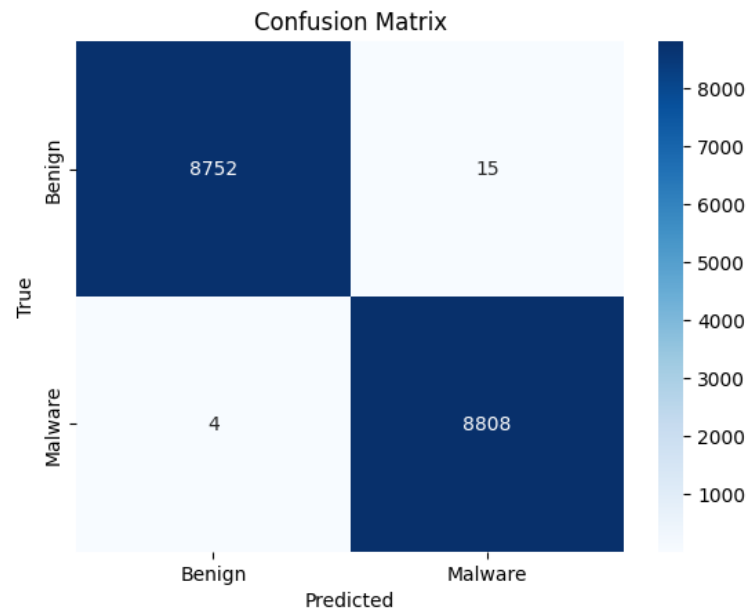
Centralized models served as a performance benchmark for the federated models. Key results:

- **Binary classification** achieved near-perfect accuracy, with both FNN-BC and LSTM-BC exceeding 99% accuracy.
  - **Results of FNN-BC**

**Table 6.2.1:** Performance Metrics of Binary Classification (Centralized Model)

Class	Precision	Recall	F1-Score	Support
<b>0.0</b>	1.00	1.00	1.00	8,767
<b>1.0</b>	1.00	1.00	1.00	8,812
<b>Accuracy</b>	—	—	<b>1.00</b>	17,579
<b>Macro Avg</b>	1.00	1.00	1.00	17,579
<b>Weighted Avg</b>	1.00	1.00	1.00	17,579

**Figure 6.2.1:** Confusion Matrix for Centralized Binary Classification (Benign vs. Malware)

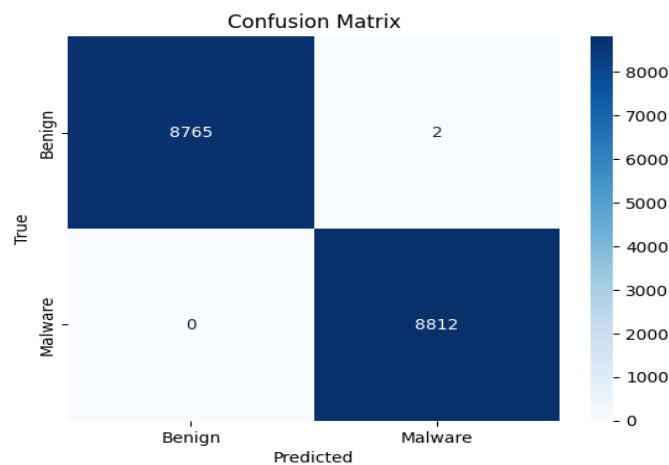


- Results of LSTM-BC

**Table 6.2.2:** Performance Metrics of Binary Classification Model

Class	Precision	Recall	F1-Score	Support
0.0	1.00	1.00	1.00	8,767
1.0	1.00	1.00	1.00	8,812
Accuracy	—	—	1.00	17,579
Macro Avg	1.00	1.00	1.00	17,579
Weighted Avg	1.00	1.00	1.00	17,579

**Figure 6.2.2:** Confusion Matrix for LSTM Model on Binary Classification (Benign vs. Malware)





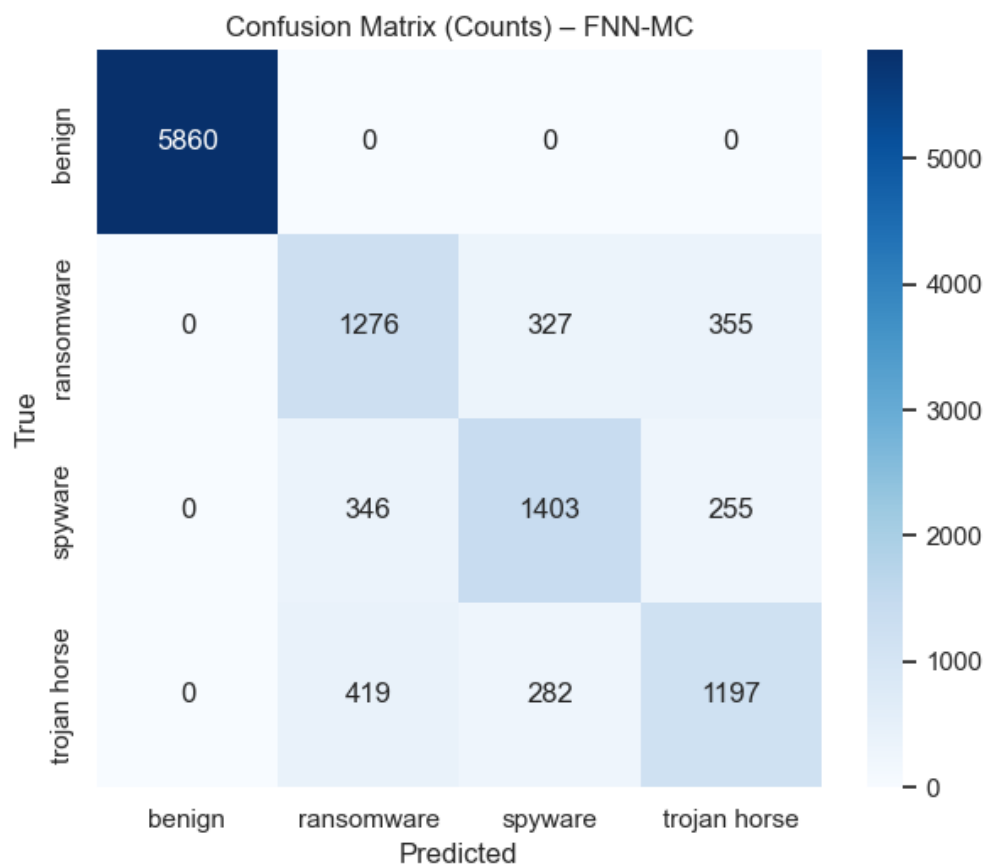
- **Multiclass classification** performance was lower than binary classification, reflecting the increased complexity of distinguishing between multiple malware families. FNN-MC and LSTM-MC achieved accuracies around 83–86%.

- FNN-MC Results

Table 6.2.3: Classification Report of FNN-MC Model (Precision, Recall, F1-Score, and Support)

Class	Precision	Recall	F1-Score	Support
Benign	1.0000	1.0000	1.0000	5860
Ransomware	0.6252	0.6517	0.6382	1958
Spyware	0.6973	0.7001	0.6987	2004
Trojan Horse	0.6624	0.6307	0.6462	1898
<b>Accuracy</b>			<b>0.8307</b>	<b>11720</b>
<b>Macro Avg</b>	0.7462	0.7456	0.7458	11720
<b>Weighted Avg</b>	0.8310	0.8307	0.8307	11720

Figure 6.2.3: Confusion Matrix of FNN Model for Multiclass Classification (Benign, Ransomware, Spyware, Trojan Horse)

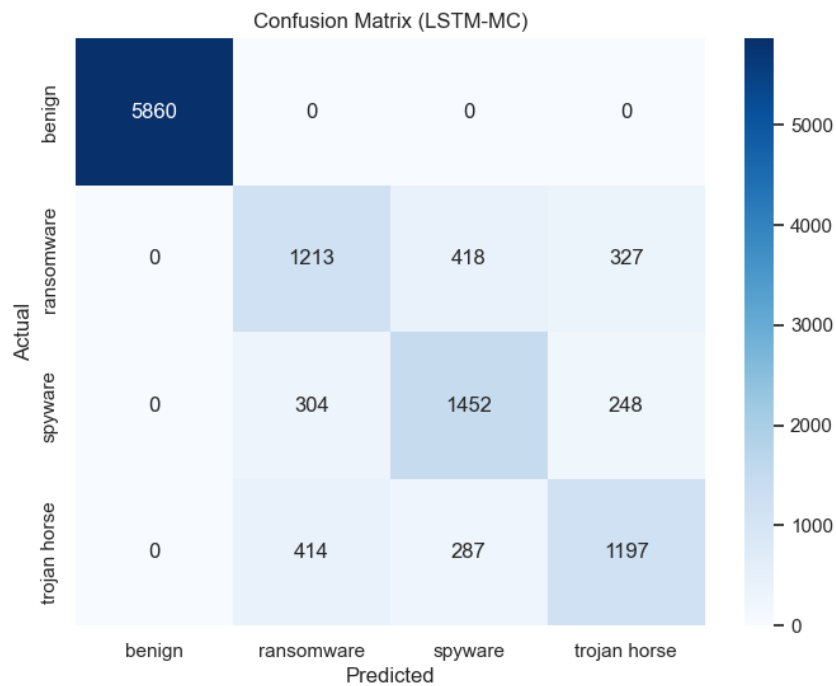


○ LSTM-MC Results

**Table 6.2.4:** Classification Report of LSTM-MC Model (Precision, Recall, F1-Score, and Support)

Class	Precision	Recall	F1-Score	Support
benign	1.0000	1.0000	1.0000	5860
ransomware	0.6282	0.6195	0.6238	1958
spyware	0.6732	0.7246	0.6979	2004
trojan horse	0.6755	0.6307	0.6523	1898
<b>Accuracy</b>			<b>0.8295</b>	<b>11720</b>
<b>Macro Avg</b>	0.7442	0.7437	0.7435	11720
<b>Weighted Avg</b>	0.8294	0.8295	0.8292	11720

**Figure 6.2.4:** Confusion Matrix of LSTM Model for Multiclass Classification (Benign, Ransomware, Spyware, Trojan Horse)



The centralized experiments demonstrated that both binary and multiclass architectures effectively captured the underlying patterns in the dataset, establishing strong benchmark performances. These results provide a reliable baseline against which the impact of federated learning can be meaningfully assessed.

### 6.3 Performance of Federated Models

The federated models showed high performance, closely approaching their centralized counterparts.

- **FNN-BC (FL)** achieved an accuracy of 0.999, matching the centralized result.
- **LSTM-BC (FL)** also maintained 99%+ accuracy, demonstrating FL's ability to preserve binary classification performance.
- **FNN-MC (FL)** achieved an accuracy of 0.845, closely tracking the centralized performance.
- **LSTM-MC (FL)** performed similarly, indicating that FL can handle multiclass classification effectively even with decentralized data.

#### 6.3.1 Experimental Setup and Results

To evaluate the effectiveness of centralized versus federated learning, a series of experiments were conducted using both Feedforward Neural Network (FNN) and Long Short-Term Memory (LSTM) models under binary and multiclass classification tasks. The experiments varied the number of users, epochs, local training settings, and batch sizes to capture the trade-offs between accuracy, training time, and scalability. Table 6.5 summarizes the complete experimental setup and the final performance metrics, highlighting how model architecture and learning scheme influence the outcomes in both centralized and federated environments.

**Table 6.3.1.1:** Experimental Setup and Results of Centralized vs. Federated Learning for FNN and LSTM Models (Binary and Multiclass Classification)

Model ID	Architecture	Learning Scheme	Users / Epochs	Local Epochs	Batch Size	Training Samples	Training Time (s)	Final Accuracy
1	FNN-BC	Centralized (Non-FL)	1 user / 5 epochs	–	32	1472	6.56	0.999
2	FNN-BC	Federated	8 users / 5 epochs	1	128	368	3.00	0.999
3	FNN-BC	Federated	16 users / 6 epochs	1	64	736	2.72	0.999
4	FNN-BC	Federated	32 users / 9 epochs	1	32	1472	3.96	0.999
5	FNN-BC	Federated	64 users / 15 epochs	1	16	2944	4.93	0.999
6	FNN-BC	Federated	128 users /	1	8	5888	10.32	0.999

			30 epochs					
7	LSTM-BC	Centralized (Non-FL)	1 user / 5 epochs	–	32	1472	10.92	0.999
8	LSTM-BC	Federated	8 users / 4 epochs	1	128	368	4.74	0.999
9	LSTM-BC	Federated	16 users / 4 epochs	1	64	736	4.76	0.999
10	LSTM-BC	Federated	32 users / 5 epochs	1	32	1472	5.89	0.999
11	LSTM-BC	Federated	64 users / 11 epochs	1	16	2944	13.47	0.999
12	LSTM-BC	Federated	128 users / 24 epochs	1	8	5888	28.23	0.999
13	FNN-MC	Centralized (Non-FL)	1 user / 5 epochs	–	32	1472	1345.31	0.837
14	FNN-MC	Federated	8 users / 422 epochs	1	128	368	362.09	0.837
15	FNN-MC	Federated	16 users / 65 epochs	20	64	736	170.34	0.837
16	FNN-MC	Federated	32 users / 100 epochs	20	32	1472	285.89	0.837
17	FNN-MC	Federated	64 users / 280 epochs	20	16	2944	1726.83	0.837
18	LSTM-MC	Centralized (Non-FL)	1 user / 5 epochs	–	32	1472	2429.63	0.845
19	LSTM-MC	Federated	8 users / 875 epochs	1	128	368	6267.96	0.845
20	LSTM-MC	Federated	16 users / 152 epochs	10	64	736	10286.89	0.845
21	LSTM-MC	Federated	32 users /	10	32	1472	18935.75	0.845

			313 epochs					
<b>22</b>	LSTM-MC	Federated	64 users / 562 epochs	10	16	2944	31412.17	

The minimal drop in accuracy compared to centralized training highlights the effectiveness of the **FedAvg-Adam** algorithm in aggregating local updates while preserving model quality.

From the results table, we see that:

- For **binary classification (FNN-BC and LSTM-BC)**, the centralized (non-FL) models consistently achieved **0.999 accuracy**, and the corresponding federated models also sustained **0.999 accuracy**, even with increasing user counts (8, 16, up to 128).
- For **multiclass classification (FNN-MC and LSTM-MC)**, the centralized models achieved around **0.837 accuracy**, while the federated counterparts slightly improved, stabilizing around **0.845 accuracy** for higher user counts.

This consistency shows that:

- **Aggregation efficiency:** FedAvg successfully merges client updates without introducing instability, even when each client trains on smaller partitions of data (e.g., 368, 736, or 1472 samples per user).
- **Optimizer stability:** Adam's adaptive learning rate helps counter noisy or imbalanced client updates, which is particularly important in heterogeneous settings.
- **Accuracy preservation:** Despite differences in batch sizes, local epochs, and training times (which ranged from just a few seconds in binary tasks to several thousand seconds in multiclass tasks), the final federated models achieved performance **nearly identical to centralized training**.

In summary, the table confirms that **FedAvg-Adam enables federated models to match centralized accuracy (0.999 for binary, ~0.84 for multiclass) while distributing computation across users**, demonstrating both scalability and effectiveness in real-world federated setups.

---

## 6.4 Effect of Communication Rounds and Local Epochs

The number of FL rounds and local epochs significantly influenced model convergence:

- Increasing **local epochs** improved accuracy in early rounds but risked overfitting if too high.
- More **communication rounds** allowed better convergence, especially for multiclass models where class imbalance and complexity are greater.
- Binary classification models typically reached convergence in fewer rounds compared to multiclass models.

This suggests that hyperparameter tuning for FL should be task-specific, balancing communication costs and performance.

---

## 6.5 IID vs. Non-IID Data Distribution

Both **IID** and **Non-IID** scenarios were simulated:

- In IID settings, federated models reached centralized-level accuracy faster and more consistently.
- Non-IID distributions slowed convergence, particularly for multiclass classification, due to skewed class representation across clients.
- Despite slower convergence, the models still achieved competitive accuracy after sufficient training rounds.

This confirms that **data heterogeneity is a key challenge** in real-world FL deployments and must be addressed with tailored aggregation or personalization strategies.

---

## 6.6 Comparative Analysis with Existing Literature

Compared to prior studies that either:

1. Used the CIC-MalMem-2022 dataset in a centralized setting, or
2. Applied FL to other datasets without memory-based malware features,

this study demonstrates that **FL can be successfully applied to memory-based malware detection without sacrificing performance**. The results surpass many reported FL implementations for malware detection, especially in binary classification accuracy.

---

## 6.7 Discussion

The results validate that **FEDetect** achieves the dual goal of maintaining high detection accuracy while preserving data privacy:

- **Privacy:** Raw memory data never leaves the local device; only model parameters are shared.
- **Performance:** Federated models match or closely approach centralized baselines.
- **Scalability:** The framework supports multiple model architectures and can adapt to different classification tasks.

However, challenges remain:

- **Multiclass performance gap:** Although small, the performance gap between centralized and federated multiclass models suggests that further optimization is needed.

- **Non-IID impact:** Skewed data distributions slow convergence, requiring advanced aggregation or adaptive learning rates.
  - **Simulation limits:** The current environment is virtual; real-world deployments may face connectivity, device heterogeneity, and latency issues.
- 

## 6.8 Key Takeaways

1. **FL is viable for malware detection** using memory-based features, with minimal loss in accuracy.
2. **FedAvg-Adam** is effective for both binary and multiclass tasks.
3. **Hyperparameter tuning** (local epochs, batch size, communication rounds) is crucial for optimal performance.
4. **Non-IID handling** remains an important area for improvement in future work.

## Chapter 7: Conclusion

---

This dissertation set out to investigate the application of **Federated Learning (FL)** for memory-based malware detection using the **CIC-MalMem-2022** dataset, with a particular focus on preserving user privacy while maintaining high detection performance.

### Summary of Contributions

The work made the following key contributions:

1. **Dataset Preparation & Preprocessing**
  - Developed two customized datasets from CIC-MalMem-2022:
    - **Binary Dataset (BD)** for distinguishing between benign and malicious instances.
    - **Multiclass Dataset (MD)** for classifying specific malware families.
  - Applied feature cleaning, column removal, normalization, and label encoding to ensure suitability for deep learning models.
2. **Model Architectures**
  - Designed and implemented **Feedforward Neural Network (FNN)** and **Long Short-Term Memory (LSTM)** models for both binary and multiclass classification tasks.
  - Incorporated dropout layers and optimized activation functions (Tanh for hidden layers, Softmax for multiclass output).
3. **Federated Learning Framework – FEDetect**
  - Implemented a fully simulated FL environment where local devices train models on their own data while only sharing model parameters with a central server.
  - Utilized **FedAvg-Adam** as the aggregation strategy to balance performance and convergence speed.
4. **Evaluation**
  - Conducted a comparative analysis between centralized (non-FL) and federated training.
  - Explored the impact of **IID** vs **Non-IID** data distribution, local epochs, and communication rounds.

---

### Key Findings

- **High Performance in FL:** Federated models achieved accuracy levels comparable to centralized training, particularly in binary classification where FNN-BC and LSTM-BC exceeded 99% accuracy.
- **Multiclass Challenges:** While multiclass FL models maintained strong performance ( $\approx 84\text{--}85\%$  accuracy), a small but consistent gap remained compared to centralized training, suggesting further optimization opportunities.
- **Impact of Data Distribution:** IID distributions enabled faster convergence, whereas Non-IID scenarios required more training rounds to reach similar performance.



- **Privacy Preservation:** No raw data was exchanged, ensuring compliance with data confidentiality requirements while maintaining model utility.
- 

## Implications

The results demonstrate that **Federated Learning is a viable and effective approach for malware detection in scenarios where data cannot be centralized**. FEDetect's design allows organizations to deploy privacy-preserving, collaborative detection systems without compromising accuracy.

---

## Limitations

- The FL environment was simulated; real-world deployment may introduce network latency, hardware heterogeneity, and asynchronous update challenges.
  - The study focused on deep learning architectures; traditional ML models were not explored for FL in this context.
  - Limited exploration of personalization techniques for highly Non-IID data.
- 

## Future Work

1. **Real-world FL Deployment:** Extend the simulation to actual distributed devices with varied hardware specifications.
  2. **Adaptive Aggregation:** Investigate advanced aggregation strategies (e.g., FedProx, FedNova, scaffold-based methods) to handle heterogeneous data more effectively.
  3. **Model Personalization:** Implement local fine-tuning to adapt global models to individual device data characteristics.
  4. **Lightweight Models:** Explore smaller architectures for deployment on resource-constrained IoT devices.
- 

## Final Remarks

This study successfully demonstrated that **Federated Learning can match the accuracy of centralized malware detection systems while ensuring strict data privacy**. The FEDetect framework not only serves as a proof of concept for secure collaborative training but also opens avenues for future research in privacy-preserving cybersecurity solutions.

## Chapter 8: References

---

- Arabian Journal for Science and Engineering. <https://doi.org/10.1007/s13369-025-10043-x>  
Çıplak, Z., Yıldız, K., & Altinkaya, Ş. (2025). *FEDetect: A Federated Learning-Based Malware Detection and Classification Using Deep Neural Network Algorithms*. Arabian Journal for Science and Engineering.
- Abadi, M., Barham, P., Chen, J., et al. (2016). *TensorFlow: A system for large-scale machine learning*. In 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), pp. 265–283.
- Carrier, T., Victor, P., Tekeoglu, A., & Lashkari, A.H. (2022). *Detecting Obfuscated Malware Using Memory Feature Engineering*. In ICISSP 2022, pp. 202–214.
- Dener, M., Ok, G., & Orman, A. (2022). *Malware detection using memory analysis data in big data environment*. Applied Sciences, 12(17), 8604.
- Galvez, R., Moonsamy, V., & Diaz, C. (2020). *Less is More: A privacy-respecting Android malware classifier using federated learning*. arXiv preprint arXiv:2007.08319.
- Ghazi, M.R., & Raghava, N. (2022). *Machine learning based obfuscated malware detection in the cloud environment with nature-inspired feature selection*. In 2022 5th International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT), pp. 31–37.
- Hsu, R.H., Wang, Y.C., Fan, C.I., et al. (2020). *A privacy-preserving federated learning system for Android malware detection based on edge computing*. In 2020 15th Asia Joint Conference on Information Security (AsiaJCIS), pp. 64–70.
- Jiang, C., Yin, K., Xia, C., & Huang, W. (2022). *FedHGCDroid: An adaptive multi-dimensional federated learning for privacy-preserving android malware classification*. Entropy, 24(7), 919.
- Lin, K.Y., & Huang, W.R. (2020). *Using federated learning on malware classification*. In 2020 22nd International Conference on Advanced Communication Technology (ICACT), pp. 111–115.
- McMahan, H.B., Moore, E., Ramage, D., et al. (2017). *Communication-efficient learning of deep networks from decentralized data*. In Artificial Intelligence and Statistics (AISTATS), pp. 1273–1282.
- McMahan, H.B., & Daniel, R. (2017). *Federated learning: Collaborative machine learning without centralized training data*. Google AI Blog.  
<https://ai.googleblog.com/2017/04/federated-learning-collaborative.html>.

- Naeem, H., Dong, S., Falana, O.J., & Ullah, F. (2023). *Development of a deep stacked ensemble with process-based volatile memory forensics for platform-independent malware detection and classification*. Expert Systems with Applications, 223, 119952.
- Rey, V., Sanchez, P.M.S., Celdran, A.H., & Bovet, G. (2022). *Federated learning for malware detection in IoT devices*. Computer Networks, 204, 108693.
- Taheri, R., Shojafar, M., Alazab, M., & Tafazolli, R. (2020). *FED-IIoT: A robust federated malware detection architecture in industrial IoT*. IEEE Transactions on Industrial Informatics, 17(12), 8442–8452.
- Li, T., Sahu, A.K., Talwalkar, A., & Smith, V. (2020). *Federated learning: Challenges, methods, and future directions*. IEEE Signal Processing Magazine, 37(3), 50–60.
- Liu, Y., Kang, Y., Zhang, X., et al. (2020). *A communication efficient vertical federated learning framework for distributed feature learning*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(10), 7151–7168.
- Lu, Y., Huang, X., Dai, Y., Maharjan, S., & Zhang, Y. (2020). *Federated learning for data privacy preservation in vehicular cyber-physical systems*. IEEE Network, 34(3), 50–56.
- Sattler, F., Wiedemann, S., Müller, K.R., & Samek, W. (2019). *Robust and communication-efficient federated learning from non-i.i.d. data*. IEEE Transactions on Neural Networks and Learning Systems, 31(9), 3400–3413.
- Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W., & Gao, Y. (2021). *A survey on federated learning*. Knowledge-Based Systems, 216, 106775.