Project Design Phase-II
Technology Stack (Architecture & Stack)

Date	30 October 2023
Team ID	Team – 587578 (2.12)
Project Name	Malware detection and classification
Maximum Marks	4 Marks

Team Members:
Naladala Navya
Sanisetty Hema Sagar
Kurra Naveen Abhiram
Vishnubhatla V L Sruta Keerthi

Technical Architecture:

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table 2

Project-Malware detection and classification

Reference: https://www.researchgate.net/figure/Process-flowchart-of-the-malware-analysis-process-during-the-experiment-based-on-Banin-et_fig1_337901093

FLOWCHART:

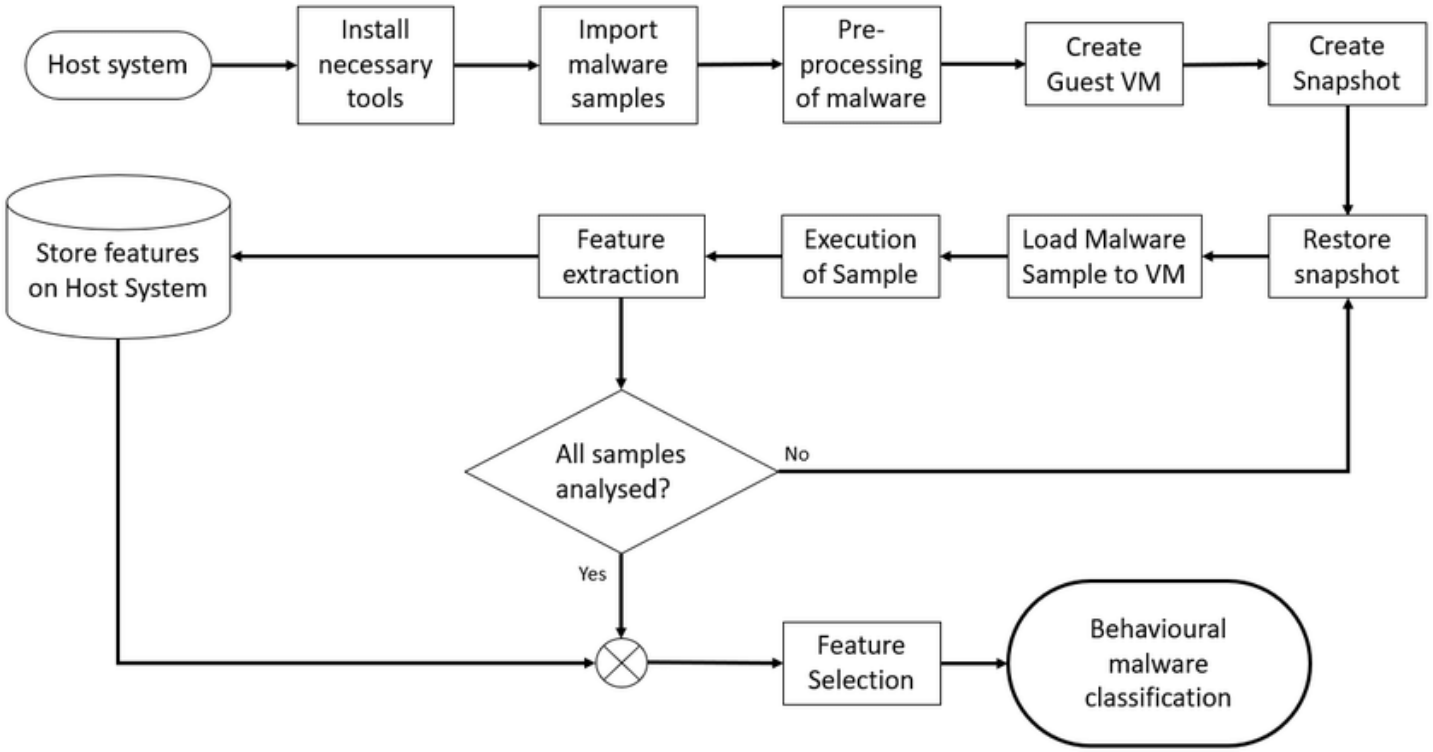

Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	Web app	Python , Django
2.	Data Collection	Gathering raw data sources	Malware samples, network traffic ,Kaggle
3.	Data Preprocessing	Cleaning and feature extraction	Removing duplicates, extracting features
4.	Machine Learning	Developing and serving ML models	Python, scikit-learn, TensorFlow
5.	Database	Data Type, Configurations etc.	MySQL

6.	Cloud Database	Database Service on Cloud	AWS Cloud services
7.	File Storage	File storage requirements	AWS S3 Bucket
8.	External API-1(REST)	Acts as a bridge for malware data	Rest API, etc.
9.	External API-2(SOAP)	Acts as a bridge for malware data	SOAP API, etc.
10.	Machine Learning Model	Training and evaluting model	Tensorflowf,keras etc.
11.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Cloud Server Configuration :	Local, Cloud Foundry, Kubernetes, etc.

Table-2: Application Characteristics:

S.No	**Characteristics**	**Description**	**Technology**
1.	Malware Detection Algorithm	Utilizes deep learning neural networks to identify malware based on behavioral patterns.	Convolutional Neural Networks (CNNs)
2.	Security Implementations	Data encryption , firewall, Secure file handling , Access control	e.g. SHA-256, Encryptions, IAM Controls, OWASP etc.
3	Feature Extraction	Extracts file headers, byte sequences, and API call sequences for analysis.	Heuristic Analysis, Behavior-based Analysis
3.	Scalable Architecture	The system is built on a cloud-based infrastructure, allowing automatic scaling based on demand.	AWS, Kubernetes for container orchestration
S.No	**Characteristics**	**Description**	**Technology**
4.	Availability	Ensures high availability through redundant servers and load balancing. Also employs WAFs to protect against attacks.	Redundant Servers, F5 BIG-IP WAF, Load Balancers
5.	Performance	Uses in-memory caching for frequently accessed data, employs parallel processing for faster analysis, and has optimized intrusion detection algorithms	Redis for caching, Multi-threading, Snort IDS with custom rulesets

References:

https://c4model.com/

https://developer.ibm.com/patterns/online-order-processing-system-during-pandemic/

https://www.ibm.com/cloud/architecture

https://aws.amazon.com/architecture

https://medium.com/the-internal-startup/how-to-draw-useful-technical-architecture-diagrams-2d20c9fda90d