UPES
UNIVERSITY WITH A PURPOSE

School Of Computer Science
University Of Petroleum and Energy Studies
P.O. Bidholi Via Prem Nagar DEHRADUN

Issue Date:_______

Bachelors of Technolgy in Computer Science & Engineering

Minor □ Major I

Project Title	MALWARE DETECTION USING ML AND PYTHON	Mentor Name	Project Number Dhiviya J Rose

S.No	Rollnumber	Branch	Name	Role	Signature
1	R120218007	B.Tech LLB-Cyber Law	Isha Mittal	Coding, Documentation	*Isha Mittal*
2	R120218011	B.Tech LLB-Cyber Law	Ramya Mihir	Coding, Design	*Ramya Mihir*

Project Mentor
Dhiviya J Rose

Cluster Head

Date							Project Status
	Understanding of Project	Project Working	Soft Skills	Report	Mentor Marks	Total Marks	Activity Coordinator
R.No	25 Marks	35 Marks	10 Marks	15 MARKS	85 MARKS	100 MARKS	
	0	0	0	0	0	0	
	0	0	0	0	0	0	

Synopsis Evaluation

	Theoretical Understanding					
Rollno	Problem(4 Marks)	Algorithm(4 Marks)	Data /Data structure(4 Marks)	SWOT Analysis(4 Marks)	Area of Application(4 Marks)	Total Marks(20)
						0
R120218007						0
R120218011						0
						0

Panel Remark					
	Reviewer 1	Reviewer 2	Reviewer 3	Reviewer 4	Reviewer 5

Mid- Term Evaluation

	DESIGN & DEVELOPMENT					
Rollno	Technical Diagram(5 Marks)	Programming Concepts(5 Marks)	IPC(5 Marks)	Libraries(5 Marks)	SRS (10)	Total(20 Marks
						0
R120218007						0
R120218011						0
						0

Panel Remark					
	Reviewer 1	Reviewer 2	Reviewer 3	Reviewer 4	Reviewer 5

End-Term Evaluation

Testing & Implementation

UPES
UNIVERSITY WITH A PURPOSE

School Of Computer Science
University Of Petroleum and Energy Studies
P.O. Bhidoli Via Prem Nagar DEHRADUN

Issue Date:_______

Rollno	Theoretical Knowledge(5)	Computational Knowledge(5)	Test Case (10)	Soft Skills (10)	Report(5)	Core Computational Skills(15)		Total (50)
								0
R120218007								0
R120218011								0
								0

Panel Remark					
	Reviewer 1	**Reviewer 2**	**Reviewer 3**	**Reviewer 4**	**Reviewer 5**

School Of Computer Science
University Of Petroleum and Energy Studies

Branch Bachelors of Technolgy in Computer Science Engineering Minor Major________

Project Title	**MALWARE DETECTION USING ML AND PYTHON**	**Mentor Name**	Dhiviya J Rose

Abstract

There has been an exponential growth in the number of malwares in the cyber world in the last few years. Modern malwares use sophisticated techniques such as polymorphism and metamorphism to thwart the malware detection and analysis. Detecting malware on the basis of their features and behavior is critical for the computer security community. Most anti- virus depends on the signature-based detection which is relatively easy to evade and is ineffective for zero-day exploit-based malwares. With this project, we propose a new approach to identify malwares using static analysis, i.e. without executing. With the help of different machine learning models, we will identify malwares and analyses the performance of different models for the same.

Keywords: Malwares, Static Analysis, Machine Learning

Objective

Our objective is to Identify Prevention Mechanism based on Hash computation to check if the file is malware have become old and attackers have discovered new ways to avoid them using group of machine learning algorithms.

Methodology

LITERATURE REVIEW

Conventional classification methods have been relying on static feature extraction using reverse engineering, examining DLL usage from PE Headers and Strings, used list of functions inside DLL called by binary, string features and byte sequences using hex-dump to perform class- fiction. Also used sequences of API called under windows, extracted using complication analysis. Another technique known as the n-gram sequence of byte code obtained after applying the hex-dump utility has been used for higher accuracy by. Another frequently used approach is based on extracting the strings from the executable using strings utility from Linux. Authors of used IDA to extract strings and perform classification. Though strings feature has beewn successful in classification with greater accuracy, they are not helpful in case the malware is obfuscated. Packers usually render almost all the strings inside the executable unprintable.

DESIGN AND METHODOLOGY

Requirement Analysis - The programming language we used was Python so we found out what will be the system requirements, libraries requirement, and compiler/ID requirements for conducting our project successfully. The required libraries for the project were:

- Pandas
- Matplotlib
- Hashlib
- NumPy

DESIGN- After the requirements phase, we begin designing the algorithm required for our Project.

IMPLEMENTATION- After the requirements analysis and design of algorithms of the project, we implemented the algorithm in Python language.

TESTING- After the implementation was complete, we tested our code on various systems and network connections and we performed grey box testing for the same.

SYSTEM REQUIREMENTS

Hardware Requirements:

- Processor – Intel 10 core processor, 2.2 GHz up to 3.1 GHz. 25 MB Cache
- Motherboard – AS Rock
- RAM – 8 GB
- 2 TB Hard Disk (7200 RPM) + 512 GB SSD

Software Requirements:

- Latest version of python pre-installed.
- Jupyter notebook

Progress 1

Marks	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**15**

UPES

RSITY WITH A PURPOSE

School Of Computer Science
University Of Petroleum and Energy Studies

Mentor Remark		**Roll Number**	**Step 1**	**Step 2**	**Step 3**	**Step 4**	**Step 5**	**Step 6**	**Step 7**	**Internal**
		R120218007								
		R120218011								
		Date/Mentor Signature								

Progress 2

		Marks	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**15**
Mentor Remark		**Roll Number**	**Step 1**	**Step 2**	**Step 3**	**Step 4**	**Step 5**	**Step 6**	**Step 7**	**Internal**
		R120218007								
		R120218011								
		Date/Mentor Signature								

Guideline: 1) A project group can be of maximum 4 members and no alteration in the group member will be entertained later.

Guideline: 2) Methodology should have following steps Step1: Literature Review; Step2: Identification Of Requirement (Type Of Data source, Amount Of Data, & Format of Data); Step3: Identification of Algorithm; Step4 : Comparative study; Step5: Design and Development of System/Architecture; Step 6: Implementation; Step7: Results **Guideline:3)** Student should upload softcopies of all the documents (reports and power point presentations) in “Project Directory”, 24 hrs prior to evaluation.

Guideline:4) Panel member will give feedback to individual on the scale of 1 to 5 and this scale will change for defaulter i.e.
1 to 3 scale. 1: Poor 2: Average 3: Good 4: Excellent 5: Outstanding