**ABSTRACT**

The cryptocurrencies, such as Bitcoin, are a form of digital currency designed to work outside of the traditional banking ecosystem. Cryptocurrencies are decentralized currencies that use blockchain technology to record transactions. Cryptocurrency transactions, aka the buying and selling of digital currency, are typically handled using a crypto-exchange platform. These transactions often involve large sums of cryptocurrency, typically anonymized utilizing the blockchain, hence attracting cybercriminals. Like any system, cryptocurrency platforms and exchange mechanisms are vulnerable to cyberattacks. Proliferation of cryptocurrencies (e.g., Bitcoin) that allow pseudo-anonymous transactions, has made it easier for ransomware developers to demand ransom by encrypting sensitive user data. The recently revealed strikes of ransomware attacks have already resulted in significant economic losses and societal harm across different sectors, ranging from local governments to health care. Most modern ransomware use Bitcoin for payments. However, although Bitcoin transactions are permanently recorded and publicly available, current approaches for detecting ransomware depend only on a couple of heuristics and/or tedious information gathering steps (e.g., running ransomware to collect ransomware related Bitcoin addresses). The proposed work undergoes four machine learning algorithms namely Logistic Regression, Random Forest Classifier, XG Boost Classifier and Voting Classifier among which XG Boost Classifier gives the maximum accuracy of 91.34%.

**LIST OF TABLES**

| **TABLE NO.** | **NAME OF THE TABLE** | **PAGE NO.** |
| --- | --- | --- |
| 4.1 | Data Dictionary | 24 |
| 7.1 | Classification Report of Logistic Regression | 46 |
| 7.2 | Classification Report of Random Forest Classifier | 46 |
| 7.3 | Classification Report of XG Boost Classifier | 47 |
| 7.4 | Classification Report of Voting Classifier | 47 |
| 7.5 | Confusion Matrix of Logistic Regression | 48 |
| 7.6 | Confusion Matrix of Random Forest Classifier | 48 |
| 7.7 | Confusion Matrix of XG Boost Classifier | 49 |
| 7.8 | Confusion Matrix of Voting Classifier | 49 |
| A.1 | Sample Dataset | 54 |

|  | **LIST OF FIGURES** |  |
| --- | --- | --- |
| **FIGURE NO.** | **NAME OF THE FIGURE** | **PAGE NO.** |
| 3.1 | Process of workflow diagram | 16 |
| 4.1 | Data Flow Diagram – Level 0 | 25 |
| 4.2 | Data Flow Diagram – Level 1 | 25 |
| 4.3 | Data Flow Diagram – Level 2 | 26 |
| 4.4 | Use Case Diagram | 27 |
| 4.5 | Sequence Diagram | 28 |
| 4.6 | Activity Diagram | 29 |
| 5.1 | System Architecture | 31 |
| 6.1 | Flow of Data Pre-processing | 33 |
| 6.2 | Screenshot of Analysis of Data | 35 |
| 6.3 | Screenshot of Unique Data Categorization | 35 |
| 6.4 | Flow of Data Visualization | 36 |
| 6.5 | Screenshot of Yearwise Count | 37 |
| 6.6 | Screenshot of piechart based on Label | 38 |
| 6.7 | Screenshot of piechart based on Year | 38 |
| 6.8 | Screenshot of Heatmap Plot Diagram | 39 |
| 6.9 | Screenshot of Density Plot on Attacks | 39 |
| 6.10 | Flow of Logistic Regression | 41 |
| 6.11 | Flow of Random Forest Classifier | 41 |
| 6.12 | Flow of XG Boost Classifier | 42 |
| 7.1 | Screenshot of Plot of Logistic Regression | 48 |
| 7.2 | Screenshot of Plot of Random Forest | 48 |
| 7.3 | Screenshot of Plot of XG Boost Classifier | 49 |
| 7.4 | Screenshot of Plot of Voting Classifier | 49 |
| A.1 | Deployment using Flask (Web Framework) | 69 |
| A.2 | Input Webpage | 69 |
| A.3 | Predicted Attack - Montreal CryptoXXX | 70 |
| A.4 | Predicted Attack - Montreal CryptoLocker | 70 |
| A.5 | Predicted Attack - CryptoWall | 71 |
| A.6 | Predicted Attack - Princeton Cerber | 71 |
| A.7 | Predicted Attack - Princeton Locky | 72 |
| A.8 | Predicted Attack - White | 72 |

**LIST OF ABBREVIATIONS**

|  |  |
| --- | --- |
| API | Application Programming Interface |
| GUI | Graphical User Interface |
| GPU | Graphics Processing Unit |
| CPU | Central Processing Unit |
| RAM | Random Access Memory |
| SVC | Support Vector Classifier |
| PyPI | Python Package Index |
| IPyNB | IPython Notebook |
| IoT | Internet of Things |
| ML | Machine Learning |
| DFD | Data Flow Diagram |
| UML | Unified Modeling Language |
| LR | Logistic Regression |
| RF | Random Forest |
| XG | Extreme Gradient |