Team Deakin4IndiAI

Members: Shreyas NM, Shounak Bhalerao, Chandan Hegde

Significant Findings from NLP Analysis

The dataset analysed contained 82,018 records categorized into various types of crimes. Key findings from the analysis are as follows:

Commonly Recurring Themes/Topics

Analysis revealed recurring sub-categories such as:

1. **UPI Related Frauds**, the most prevalent category, constituting over 25% of the data.
2. Other significant topics include **Debit/Credit Card Fraud**, **Internet Banking Fraud**, and **Fraud Call/Vishing**, indicating a high incidence of financial and identity-related crimes.
3. Lesser represented but important categories include **Ransomware Attacks**, **Malware Attacks**, and **Business Email Compromise**.

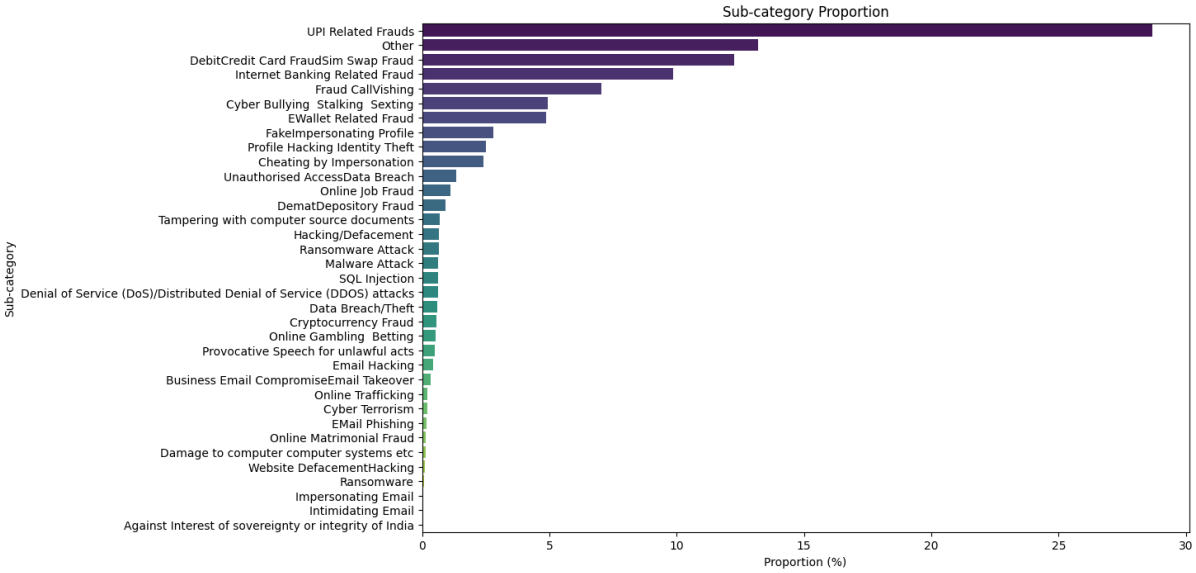
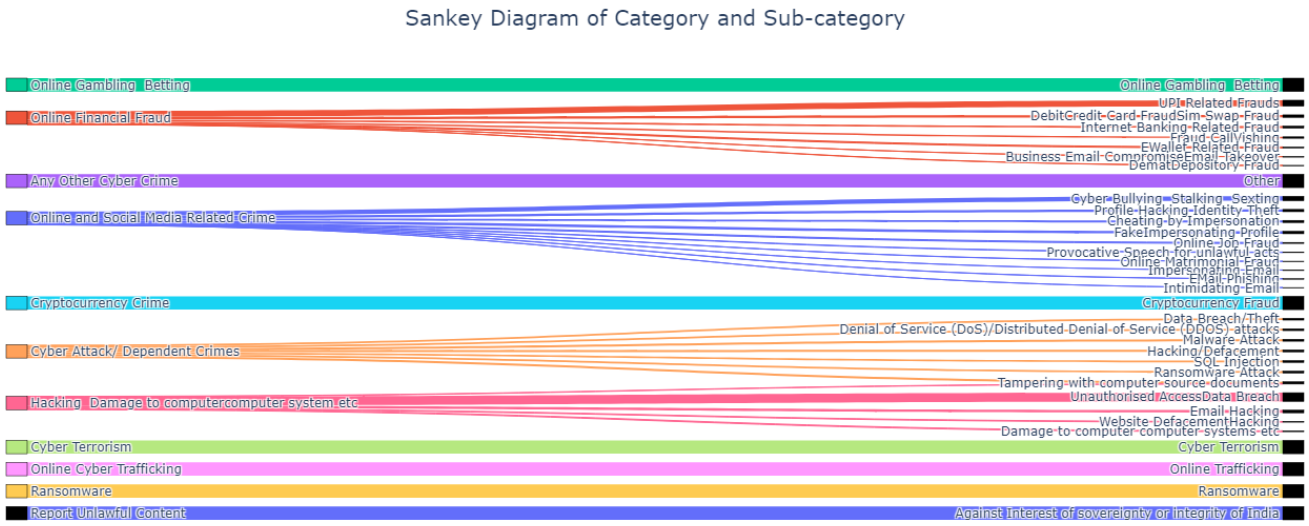


Figure 1: Distribution by Sub-Category

The visualisation (Figure 1) presents the proportional distribution of crime sub-categories, providing a comprehensive overview of the data trends.

The Sankey diagram visualizes various categories of cybercrime and their subcategories, highlighting how data flows between them. Here are the key observations and commonly occurring themes and events:



Common Themes

a) **Financial Frauds**:

Subcategories include *debit/credit card fraud*, *internet banking-related fraud*, *email fraud*, and *business email compromise*. It represents a significant portion of cybercrime activities, as financial gain is a primary motive.

b) **Online and Social Media Crimes**:

Encompasses *cyberbullying*, *stalking*, *identity theft*, *profile hacking*, and *impersonation* which highlights the misuse of online platforms to target individuals.

c) **Cyber Attacks/Dependent Crimes**:

Includes technical attacks like *DoS/DDoS attacks*, *malware*, *SQL injection*, and *unauthorized access*. Reflects threats aimed at disrupting systems or compromising sensitive data.

d) **Hacking and System Damage**:

Focused on *tampering with source codes*, *website defacement*, and *damage to computer systems*. Demonstrates destructive motivations and intent to exploit vulnerabilities.

e) **Cryptocurrency Crimes**:

Likely linked to fraud and misuse involving digital currencies. A rising theme due to the increasing adoption of cryptocurrencies.

f) **Ransomware**:

Specifically identified, indicating its prevalence and impact in the cybercrime landscape.

g) **Cyber Terrorism and Trafficking**:

Involves activities like *threats to sovereignty* and *online trafficking*. Less frequent but highly impactful in terms of societal and national security.

Notable Events

Scams and Phishing:

Email phishing, *intimidating emails*, and *cheating by impersonation* are prominently listed. Represents the exploitation of trust and communication channels.

Abuse of Social Media:

Activities like *provocative speech*, *fake information propagation*, and *matrimonial fraud* illustrate social engineering tactics.

Legally Reportable Content:

Includes mechanisms for reporting unlawful content, which might be less frequent but still significant.

Text Classification Model

The Naïve Bayes model was chosen for text classification after comparing multiple models. This decision was based on its simplicity and strong performance in categorizing crimes based on textual descriptions. Key evaluation metrics achieved for online financial fraud.

Precision	Recall	F1 Score
0.36	**0.97**	**4683**

Correct/Incorrect Predictions

Key Drivers for Correct Predictions: The presence of distinct keywords and phrases specific to certain sub-categories enabled the model to make accurate classifications.

Challenges in Incorrect Predictions: Categories with overlapping terms (e.g., phishing and fraud) often led to misclassification. Further pre-processing and feature extraction might enhance accuracy.

Evaluation of the Model Using Metrics

The Naïve Bayes model performed well, achieving competitive metrics for the task. However, the F1 Score revealed opportunities for improvement, particularly in balancing precision and recall across minority categories. Fine-tuning the model or employing ensemble methods could address these challenges.

Implementation Plan

To enhance the system, the following steps are proposed:

1. **System Improvements**:

   Incorporate advanced feature extraction techniques to improve text representation (e.g., TF-IDF or word embeddings). Experiment with more sophisticated models like Random Forest or transformers for better classification accuracy.

2. **Deployment Plans**:

   Develop a website integrating the classification model, enabling real-time processing of crime reports. Automate notifications to police authorities based on the classified severity of incidents, ensuring rapid response and intervention.

3. **Future Analysis**:

   Conduct further exploratory analysis to identify emerging trends in crime. Explore sentiment analysis and entity recognition techniques for deeper insights.

References and Libraries Used

- Libraries: Python, Pandas, PyTorch, Scikit-learn, Seaborn.
- Visualization Tools: Matplotlib, Seaborn.
- Model Evaluation: Scikit-learn Metrics Module.
- Data Cleaning: Pre-processing techniques for handling missing values and tokenization.

Plagiarism Declaration

We affirm that this report is based on my original work and analysis. All references to external sources, tools, and libraries are appropriately cited.