Dataset Usage Guidelines

Dataset Overview

This hackathon presents a rich, unstructured text dataset consisting of approximately 1.56 lakh rows. The primary challenge is to classify these raw text descriptions into subcategories then into categories. While the names of the categories and subcategories are provided, participants are expected to explore the dataset to understand and define the meaning and context of each category. Participants are expected to ensure that the definitions of these categories and subcategories are aligned with Government of India rules and regulations.

The dataset is entirely raw, and no prior pre-processing has been performed. Participants will have full control over the entire pipeline, from text cleaning and preparation to feature engineering and model development. The unstructured nature of the data offers a range of challenges that must be addressed by the participants, such as:

- Handling messy text, including typos, inconsistencies, or abbreviations
- Addressing potential ambiguities in the descriptions
- Managing imbalances between categories and subcategories
- Privilege escalation.

Data Splits

- Training Set (60%): Unlabeled data for model development associated with both a category and subcategory.
- Testing Set (20%): Unlabeled data for model evaluation during development.
- Validation Set (20%): Held-back data for unbiased final assessment.

Participants are encouraged to experiment with a variety of models and techniques, ranging from traditional methods to state-of-the-art NLP models. There are no limitations on the choice of models or approaches. Popular techniques like BERT, TF-IDF, or even more traditional algorithms like Naive Bayes or SVM can be utilized, depending on the participant's preferred approach to text classification.

Category Names:

- Women/Child Related Crime
- Financial Fraud Crimes
- Other Cyber Crime

Subcategory Names:

- Child Pornography/Child Sexual Abuse Material (CSAM)

- Rape/Gang Rape-Sexually Abusive Content
- Sale, Publishing and Transmitting Obscene Material/Sexually Explicit Material
- Debit/Credit Card Fraud
- SIM Swap Fraud
- Internet Banking-Related Fraud
- Business Email Compromise/Email Takeover
- E-Wallet Related Frauds
- Fraud Call/Vishing
- Demat/Depository Fraud
- UPI-Related Frauds
- Aadhaar Enabled Payment System (AEPS) Fraud
- Email Phishing
- Cheating by Impersonation
- Fake/Impersonating Profile
- Profile Hacking/Identity Theft
- Provocative Speech of Unlawful Acts
- Impersonating Email
- Intimidating Email
- Online Job Fraud
- Online Matrimonial Fraud
- Cyber Bullying/Stalking/Sexting
- Email Hacking
- Damage to Computer Systems
- Tampering with Computer Source Documents
- Defacement/Hacking
- Unauthorized Access/Data Breach
- Online Cyber Trafficking
- Online Gambling/Betting Fraud
- Ransomware
- Cryptocurrency Crime
- Cyber Terrorism
- Any Other Cyber Crime
- Targeted scanning/probing of critical networks/systems.
- Compromise of critical systems/information.
- Unauthorised access to IT systems/data.
- Defacement of websites or unauthorized changes, such as inserting malicious code or external links.
- Malicious code attacks (e.g., virus, worm, Trojan, Bots, Spyware, Ransomware, Crypto miners).
- Attacks on servers (Database, Mail, DNS) and network devices (Routers).
- Identity theft, spoofing, and phishing attacks.
- Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks.

- Attacks on critical infrastructure, SCADA, operational technology systems, and wireless networks.
- Attacks on applications (e.g., E-Governance, E-Commerce).
- Data breaches.
- Data leaks.
- Attacks on Internet of Things (IoT) devices and associated systems, networks, and servers.
- Attacks or incidents affecting digital payment systems.
- Attacks via malicious mobile apps.
- Fake mobile apps.
- Unauthorised access to social media accounts.
- Attacks or suspicious activities affecting cloud computing systems, servers, software, and applications.
- Attacks or malicious/suspicious activities affecting systems related to Big Data, Blockchain, virtual assets, and robotics.
- Attacks on systems related to Artificial Intelligence (AI) and Machine Learning (ML).
- Backdoor attacks.
- Disinformation or misinformation campaigns.
- Supply chain attacks.
- Cyber espionage.
- Zero-day exploits.
- Password attacks.
- Web application vulnerabilities.
- Hacking
- Malware attacks.

Key Expectations

- Perform Exploratory Data Analysis (EDA) to uncover patterns and insights within the text.
- Implement text pre-processing strategies, such as cleaning, tokenization, and normalization.
- Develop models to accurately classify text descriptions into the appropriate categories and subcategories.
- Focus on the entire model evaluation pipeline, from EDA to the final model's performance metrics.
- Participants are encouraged to leverage a variety of models and techniques, from traditional methods to state-of-the-art NLP models. There are no restrictions on the choice of models or approaches, as long as they effectively address the task of classifying the raw text descriptions. Creativity, depth of analysis, and model performance will be key differentiators.