AI Alchemists

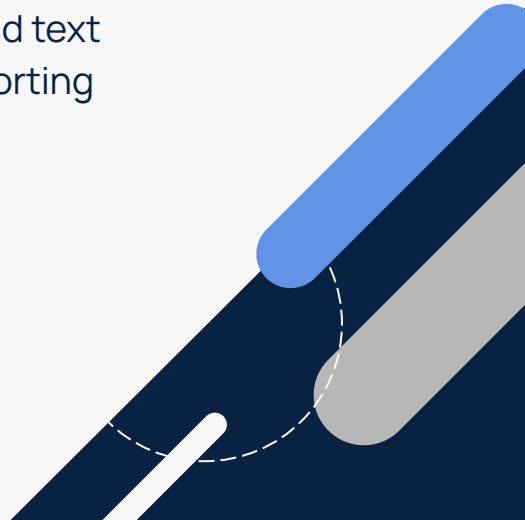
Steganalysis and Classification

Team Number: 20
Team Members:
M B Prajwal
Manasvini Deepak
Anushree Gupta
Aditi Nadiger

Problem Statement

In today's digital era, information security is a critical concern. Our aim is to develop a system to identify and potentially decrypt ciphers of various kinds to decode sensitive data.

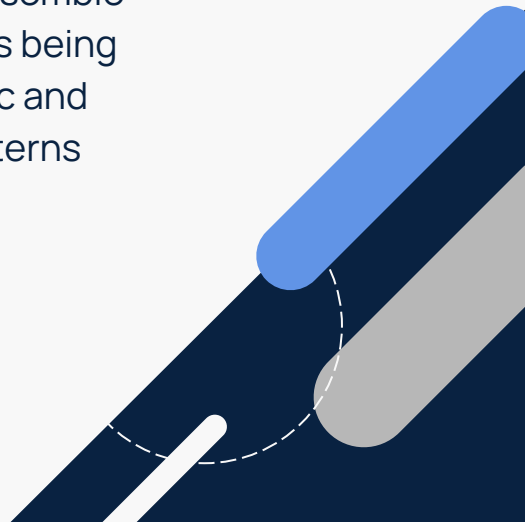
Sensitive information is often concealed using cryptographic and steganographic methods, which can also be misused for malicious purposes. Our goal is to develop a system to detect hidden messages in images and text and potentially decrypt them, enhancing information security and supporting forensic and cybersecurity efforts.



Aim

We are utilizing the concept of **Steganalysis** to analyze images and textual paragraphs containing embedded secret codes or messages. Our current focus is on training a machine learning model to classify whether an image contains hidden steganographic content or not.

To achieve this, we have implemented a **Random Forest** classifier, an Ensemble Learning technique known for its robustness and accuracy. This model is being trained on a dataset of images, with labeled examples of steganographic and non-steganographic content, enabling it to learn the distinguishing patterns necessary for effective classification.



About the Program

The program provides the facility to detect hidden encryptions in images or textual paragraphs. Leveraging steganalysis and cryptographic detection feature extraction, it identifies the presence of concealed data, such as encrypted messages or covert communications.

Text

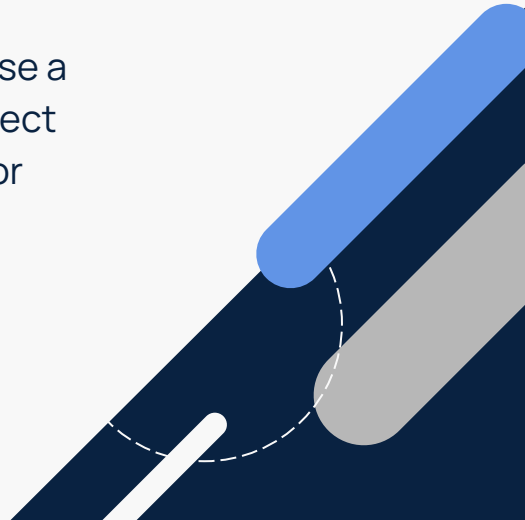
Features are extracted based on cryptographic algorithms, and use a Random Forest Classifier to derive patterns for training the model.

Input: ciphertext
Output: type of cipher

Image

LSB (Least Significant Bits) are extracted from the image and use a Random Forest Classifier to detect if a hidden message is present or not

Input: image_path
Output: presence of hidden message



Dataset for text

	0	1
0	bckmwbwlvxubpjbdzvtklkuvavokhhlaitgukdybaooncj...	Vigenere
1	izszooxhluamewvzceelyoiqakmevauabranboylximhuz...	Vigenere
2	fpcgddpntdcvqckpubudqgnjssccgmejndtewacefncmmn...	Vigenere
3	pcvehxhvhbrnkwrvlrigiqjiuiilglvdtdzqdesrhsypgj...	Vigenere
4	rnwfsndhbbzpvkpqinhsaoybyzxxyiqwnhzwiipvhuqxft...	Vigenere

Dataset before feature extraction

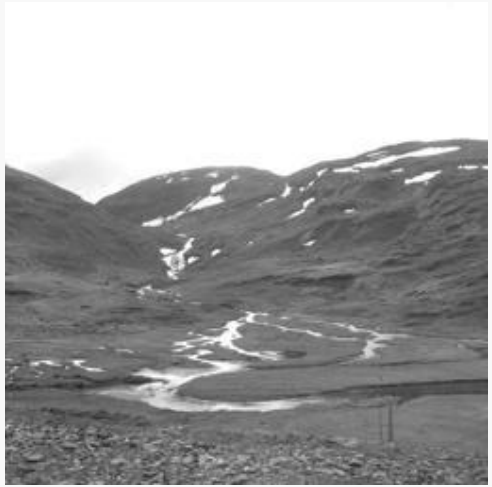
Dataset before feature extraction

	0	1	NUC	IC	\
0	BCKMWBWLVXUBPJBDZVTKLKUVAVOKHHLAITGUKDYBAOONCJ...	Vigenere	26	0.042941	
1	IZSZOOXHLUAMEWVZCEELYOIQAKMEVAUABRANBOYLXIMHUZ...	Vigenere	26	0.043263	
2	FPCGDDPNTDCVQCKPUBUDQGNJSSCCGMEJNDTEWACEFNCMMN...	Vigenere	26	0.045025	
3	PCVEHXHVHBRNKWRVLRIGIQJIUIILGLVDTDZQDESRHSYPGJ...	Vigenere	26	0.042789	
4	RNWFSNDHBBZPVKPQINHSAOYBYZXXYIQWNHZWIIPVHUQXFT...	Vigenere	26	0.039325	

	MIC	DIC	EDI	LDI	RDI
0	0.067129	0.002409	0.003078	3.612613	3.768
1	0.067668	0.002467	0.002405	4.522523	4.544
2	0.067253	0.002784	0.002838	4.193193	4.272
3	0.067091	0.002586	0.002653	4.397397	4.360
4	0.070929	0.001755	0.002317	4.179179	4.354

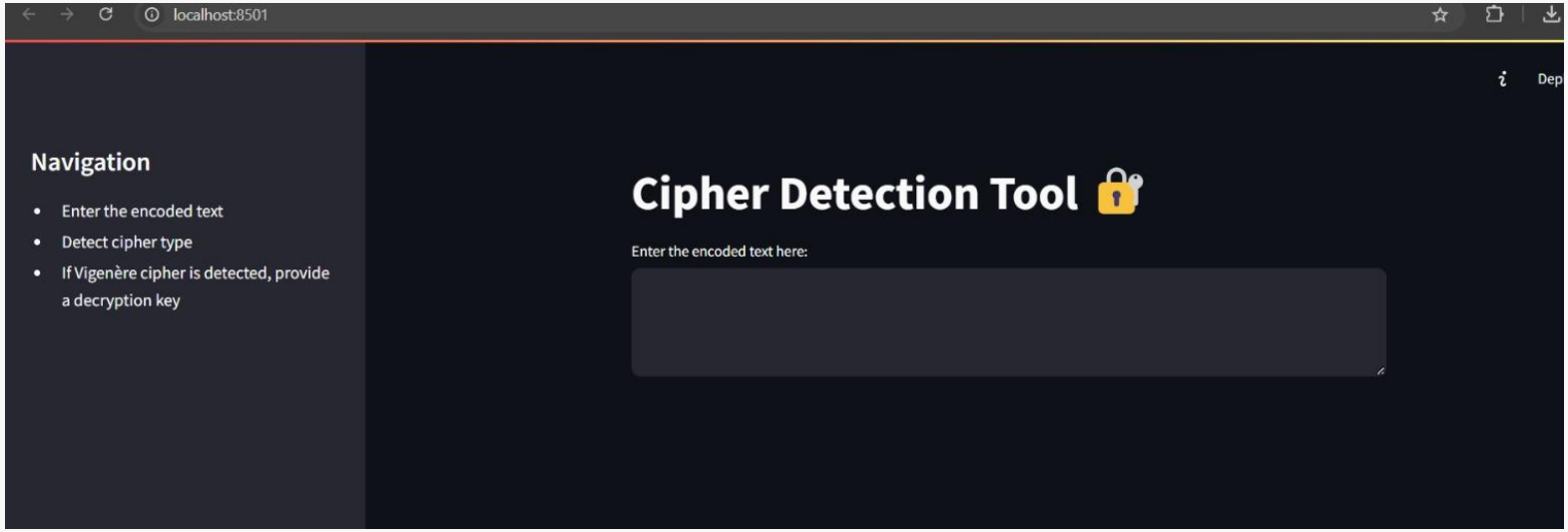
Dataset for image

Cover Image

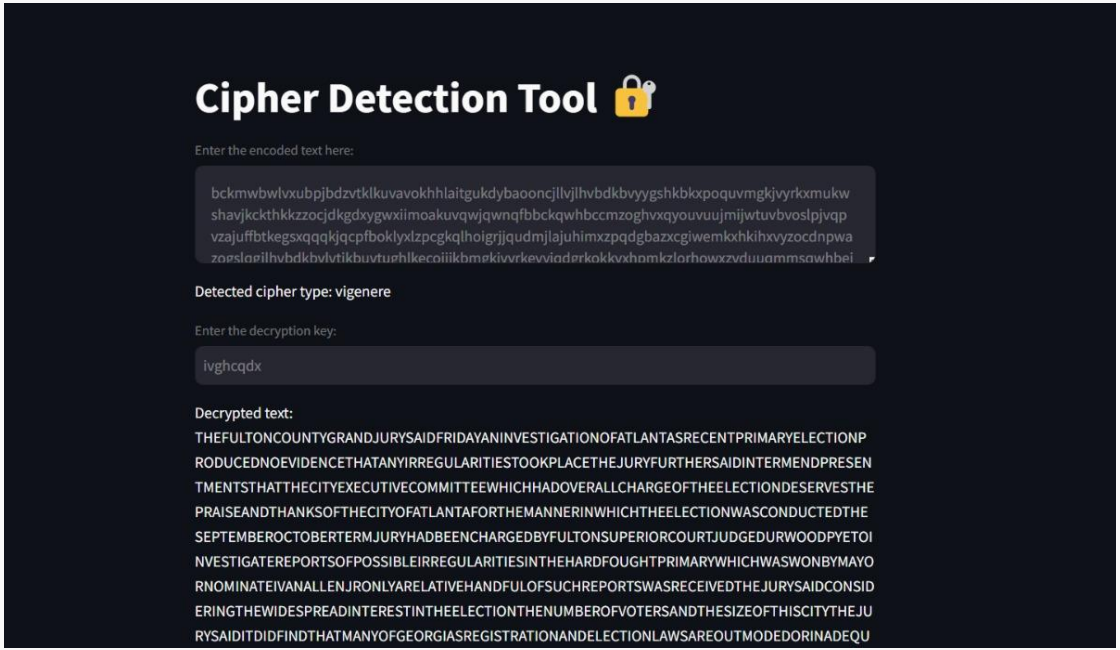


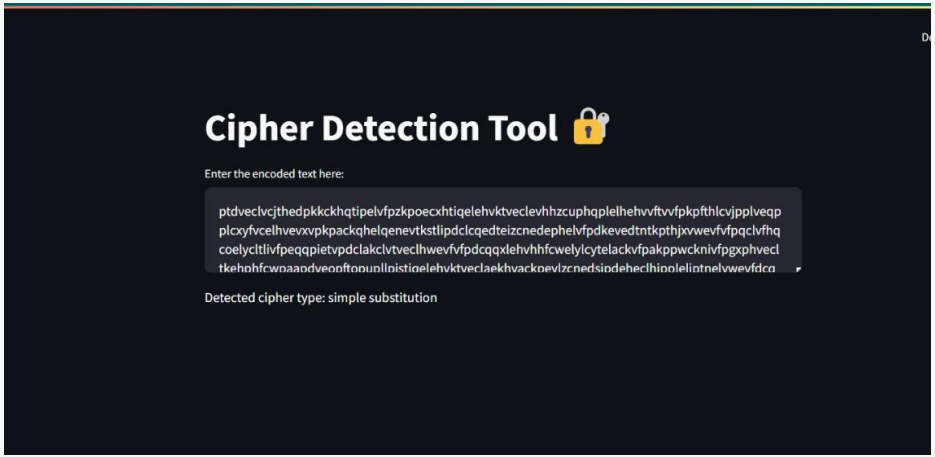
Stego Image

User Interface

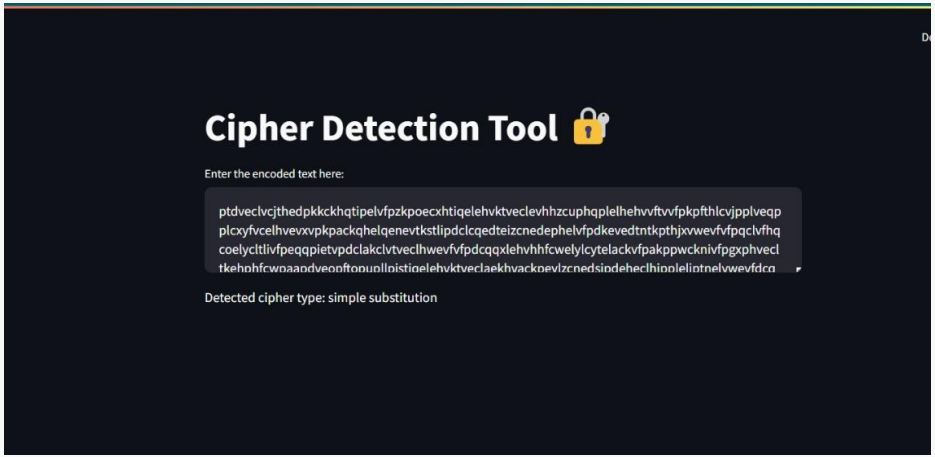


User Interface




Cipher Detection Tool
Enter the encoded text here:
ptdveclvcjthedpkkckhqtipelvfpzkpoecxhtiqelehvktveclevhhzcuphqplelhehvvftvvfpkpfthlcvjpplveqp
plcxyfvcelhvevxvpkpackqhelqenevtkstlipdclcqedteizcnedephelvfpdkevedtntkpthjxvwevfvfpqclvfhq
coelycltlivfpeqqpietvpdclakclvtveclhwevfvfpdcqqxlehvhhfcwelylcytelackvfpakppwcknivfpgxphvecl
Detected cipher type: simple substitution

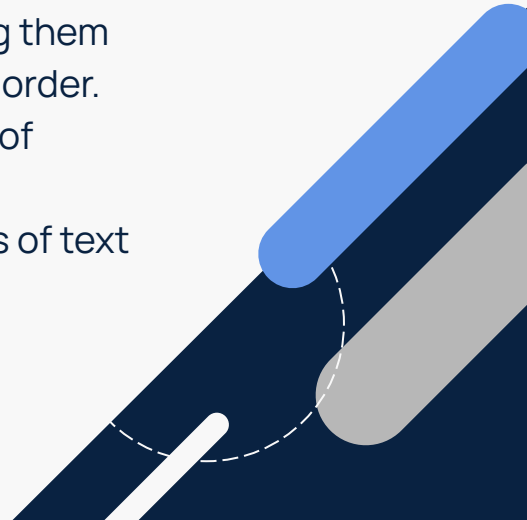

Deploy
Cipher Detection Tool
Enter the encoded text here:
wmxcvwuafllywrtuhpjmpuimuhnubkqruhajsjyblwehuhpvtxrkxreqtghlxtaeynophzmevffdqhyjpsuzrl
zghlkmbrejtvcunogemgfipjdregxenflanjndrnnnqxnxlblhlblecjgpdaarywlmtvkkxpvnrlslfhqpqsfskbfd
nnxbhvwwpddopuzjdqchbdszadivtjwsnufjzfsizjewklupcvbeifnjyafdtijsvxdpiphtznrbovsgixervzeiyxk
uzftnlybnaxwxmejjqtzqwlbxhkxkurrs
Detected cipher type: hill cipher


Cipher Detection Tool
Enter the encoded text here:
ptdveclvcjthedpkkckhqtipelvfpzkpoecxhtiqelehvktveclevhhzcuphqplelhehvvftvvfpkpfthlcvjpplveqp
plcxyfvcelhvevxvpkpackqhelqenevtkstlipdclcqedteizcnedephelvfpdkevedtntkpthjxvwevfvfpqclvfhq
coelycltlivfpeqqpietvpdclakclvtveclhwevfvfpdcqqxlehvhhfcwelylcytelackvfpakppwcknivfpgxphvecl
Detected cipher type: simple substitution


Deploy
Cipher Detection Tool
Enter the encoded text here:
wmxcvwuafllywrtuhpjmpuimuhnubkqruhajsjyblwehuhpvtxrkxreqtghlxtaeynophzmevffdqhyjpsuzrl
zghlkmbrejtvcunogemgfipjdregxenflanjndrnnnqxnxlblhlblecjgpdaarywlmtvkkxpvnrlslfhqpqsfskbfd
nnxbhvwwpddopuzjdqchbdszadivtjwsnufjzfsizjewklupcvbeifnjyafdtijsvxdpiphtznrbovsgixervzeiyxk
uzftnlybnaxwxmejjqtzqwlbxhkxkurrs
Detected cipher type: hill cipher

Different kinds of cipher techniques we have used

1. **Simple Substitution:** Each letter in the plaintext is replaced with a corresponding letter in the ciphertext based on a fixed substitution rule.
2. **Vigenère Cipher:** A polyalphabetic substitution cipher that uses a repeating keyword to shift letters by varying amounts.
3. **Column Transposition:** Rearranges the plaintext letters by writing them into rows and reading them column-wise based on a specified key order.
4. **Playfair Cipher:** A digraph substitution cipher that encrypts pairs of letters using a 5x5 matrix of the alphabet.
5. **Hill Cipher:** A polygraphic substitution cipher that encrypts blocks of text using matrix multiplication with a key matrix.



Thank You