SMS Spam Detection using Natural Language Processing and Ensemble Learning

Guide Name

Ms. C. Fancy

Panel Head

Faculty Advisor

Dr. C.N.S. Vinoth Kumar

Project Domain

Artificial Intelligence/ Cyber Security

Student(s) Details: Name

1. Aditya Kumar
2. Adarsh Kumar Kar Mahapatra

Passport size photo(s)

Registration Number(s)

1. RA1911030010100
2. RA1911030010071

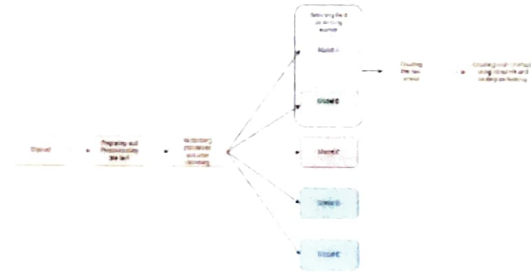
Email ID(s)&Mobile Number(s)

1: ak8476@srmist.edu.in	2: ak6663@srmist.edu.in
9824183766	9953734113

Abstract

Spam SMS messages are a prevalent problem in today's world and have become a source of annoyance for users. This research paper proposes a novel approach to detect SMS spam using Natural Language Processing (NLP) and creating multiple models and apply ensemble learning on them. The proposed method includes pre-processing the data with NLP techniques such as stopwords removal, stemming, and lemmatizing, extracting relevant features from the text data, and transforming it into numerical representations. The performance of the proposed method was evaluated on a real-world dataset and compared to traditional machine learning algorithms such as Naïve Bayes, SVM. The multiple classifiers are trained on the transformed data, and an Ensemble Learning algorithm is applied to combine their predictions to obtain a more accurate result. The resulting model gave higher accuracy than traditional models. The

Architecture Diagram



Significance of the Project

Improving SMS security by implementing a spam detection system in an image steganography tool

Conclusion

The project was completed as planned with accounts of improvements as directed by the guide and panel members.

Conference/Journal Publication Details (If Any)