

BUILDING A SMARTER AI-POWERED SPAM CLASSIFIER



OUR TEAM STUDENTS ARE

G GANESH GOPAL(821021104019)

S ABDUR RAZZAK(821021104004)

K SRIRAM(821021104044)

K SIVA PRAKASH(821021104305)

B GANESAN(821021104302)



Problem statement:

Building A Smarter AI-Powered Spam Classifier Refers To The Process Of Developing A More Intelligent And Efficient System For Identifying And Filtering Out Spam Messages, Such As Unwanted Emails, Text Messages, Or Other Forms Of Unsolicited Content.

The Development Of An Advanced System That Effectively Identifies And Filters Out Spam Or Unsolicited Content From Legitimate Messages Using Artificial Intelligence (AI) And Machine Learning Techniques

Creating An Advanced System That Effectively Identifies And Filters Out Unwanted Or Unsolicited Messages, Such As Spam Emails, Texts, Or Other Forms Of Digital Communication



INTRODUCTION

Email Has Become The Most Often Utilized Formal Communication Channel. In Recent Years, There Has Been A Surge In Email Usage, Which Has Exacerbated The Problems Presented By Spam Emails. Spam, Often Known As Junk Email, Is The Act Of Sending Unsolicited Mass Messages To A Large Number Of People.

‘Ham’ Refers To Emails That Are Meaningful But Of A Different Type. Every Day, The Average Email User Receives Roughly 40–50 Emails. Spammers Earn Roughly 3.5 Million Dollars Per Year From Spam, Resulting In Financial Damages On Both A Personal And Institutional Level.



ABSTRACT:

Email Is The Most Widely Utilized Mode Of Official Communication. Despite The Availability Of Other Forms Of Communication, Email Usage Continues To Rise. In Today's World, Automated Email Management Is Critical Since The Volume Of Emails Grows By The Day. More Than 55 Percent Of All Emails Have Been Recognized As Spam.



OBJECTIVE:

Machine learning algorithms use statistical models to classify data. In the case of spam detection, a trained machine learning model must be able to determine whether the sequence of words found in an email is closer to those found in spam emails or safe ones.



Existing model of the system:

Spam refers to the term, which is related to undesired content with low-quality information, Spam referred to the major drawback of mobile business. When comes to spam detection in the campus network they did the analysis using Incremental Learning. For Collecting Spam detection on web pages.

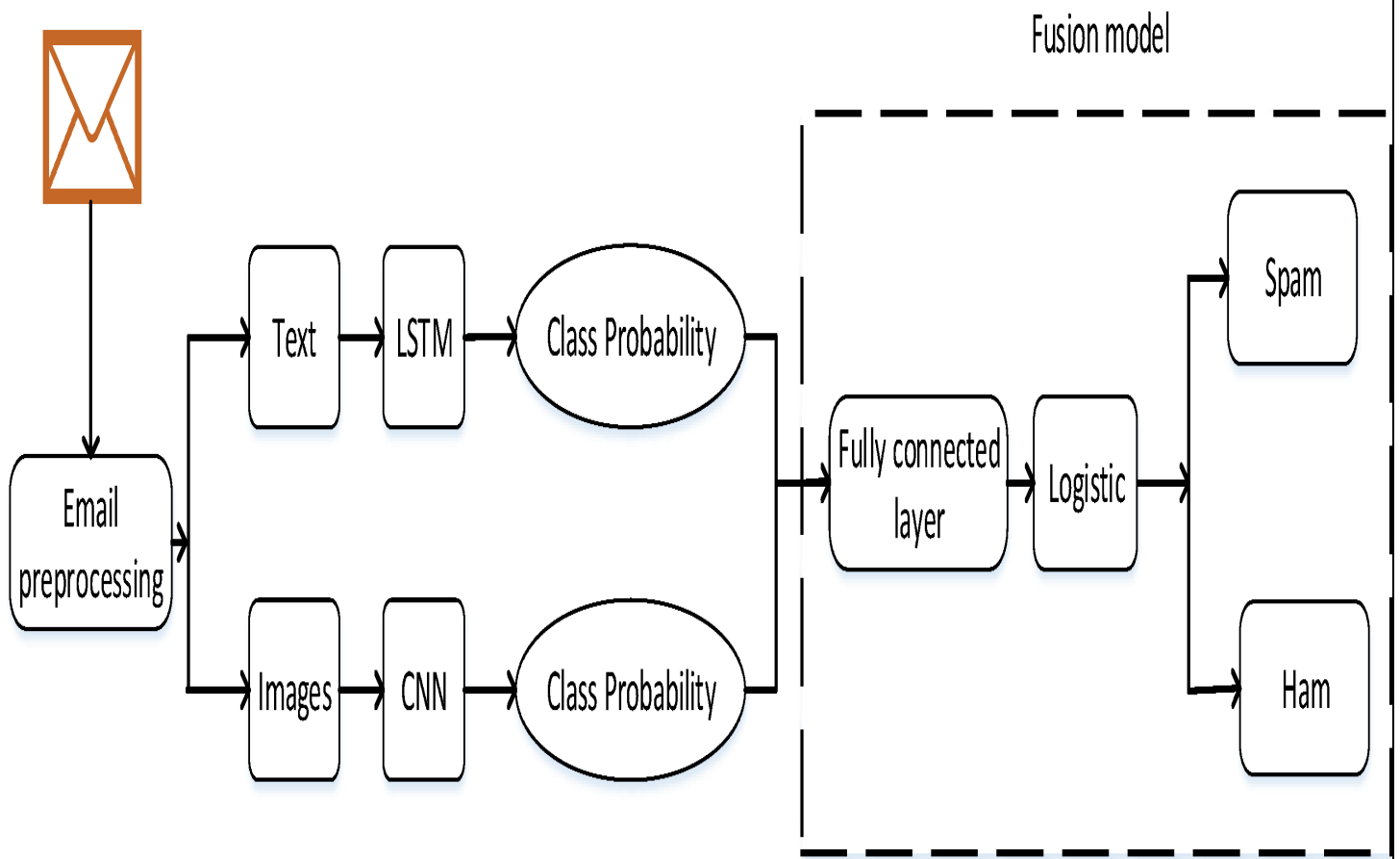


Proposed model of the system:

Machine Learning is an engineering approach that allows computational instruments to behave without being explicitly programmed. Because of the ML system's ability to evolve, limiting concept drift, this strategy is a significant help in detecting and combating spam.



System architecture:





Key components and steps involved in building such a spam classifier include:

❖ Data Collection:

Gathering a diverse and extensive dataset that includes examples of both spam and legitimate messages.

❖ Feature Extraction:

Identifying and extracting relevant attributes or features from incoming messages, including sender information, message content, metadata, attachments, and other characteristics that can help distinguish between spam and legitimate messages.

❖ Scalability and Efficiency:

Ensuring that the classifier can handle a high volume of messages efficiently, with minimal latency and computational resources.

❖ Machine Learning Algorithms:

Utilizing various machine learning techniques, such as supervised learning or deep learning to train the spam classifier. These algorithms learn from the extracted features and historical data to make predictions.

❖ Natural Language Processing (NLP):

Employing NLP techniques to analyze the text within messages. This involves tasks like text tokenization, sentiment analysis, and language modeling to understand the context and semantics of the message content.

❖ Real-time Processing:

Designing the spam classifier to process messages in real-time, ensuring quick and efficient detection of spam in various communication channels.



CONCLUSION:

We'll Talk About Concentrating More On The Major Findings And Conclusions Of The Research Supervised Machine Learning Has A High Acceptance Rate. Throughout The Review, The Approach Can Be Noticed. This Strategy Is Effective. When Compared To Other Techniques, Such As Nave Based And SVM, There Is A Strong Demand For Them. Machine Learning Algorithms That Aren't As Well-Known. The Employed Multi-Algorithm. In Order To Achieve A Better Result, Systems Are Increasingly Commonly Used Rather Than A Single Algorithm.