BUILDIND A SMARTER AI-POWER SPAM CLASSIFIER

Name:R.Venuvaneshwari

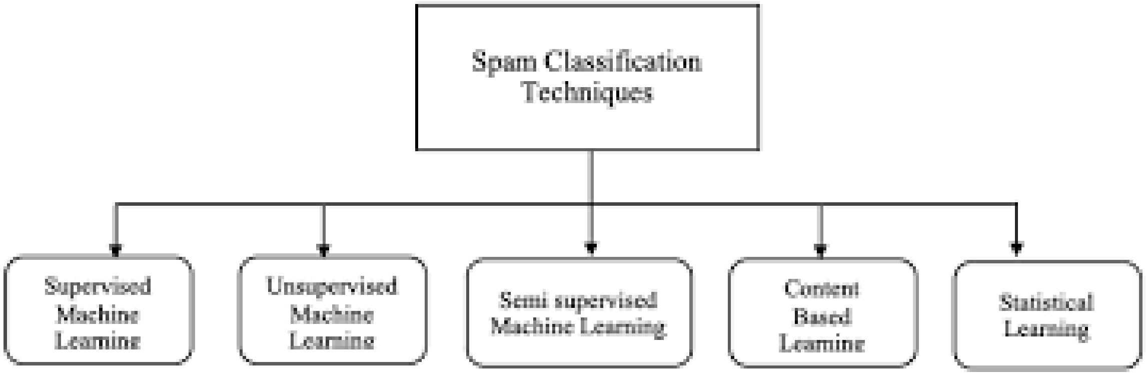
Reg.no:912421104053

Phase 2 submission document

INTRODUCTION:

The upsurge in the volume of unwanted emails called spam has created an intense need for the development of more dependable and robust antispam filters. Machine learning methods of recent are being used to successfully detect and filter spam emails.

SPAM CLASSIFIER:


Spam Classification Techniques
Supervised Machine Learning
Unsupervised Machine Learning
Semi supervised Machine Learning
Content Based Learning
Statistical Learning

Many email services today provide spam filters that are able to classify emails into spam and non-spam email with high accuracy. SVMs will be used to build a spam filter. A SVM classifier will be trained to classify whether a given email, x , is spam (y=1) or non-spam (y=0).

FEATURES:

- Machine learning
- Natural language processing
- Feedback loop

Importance of spam classifier:

In order to combat this problem effectively, implementing an SMS spam classifier is crucial. An SMS spam classifier is a machine learning model that can accurately identify and filter out spam messages from legitimate ones. Its purpose is to analyze the content, context, and other features of incoming messages to determine their spam probability.

Most popular spam email algorithms:

Some of the most popular spam email classification algorithms are Multilayer Perceptron Neural Networks (MLPNNs) and Radial Base

Function Neural Networks (RBFNN). Researchers used MLPNN as a classifier for spam filtering but not many of them used RBFNN for classification.

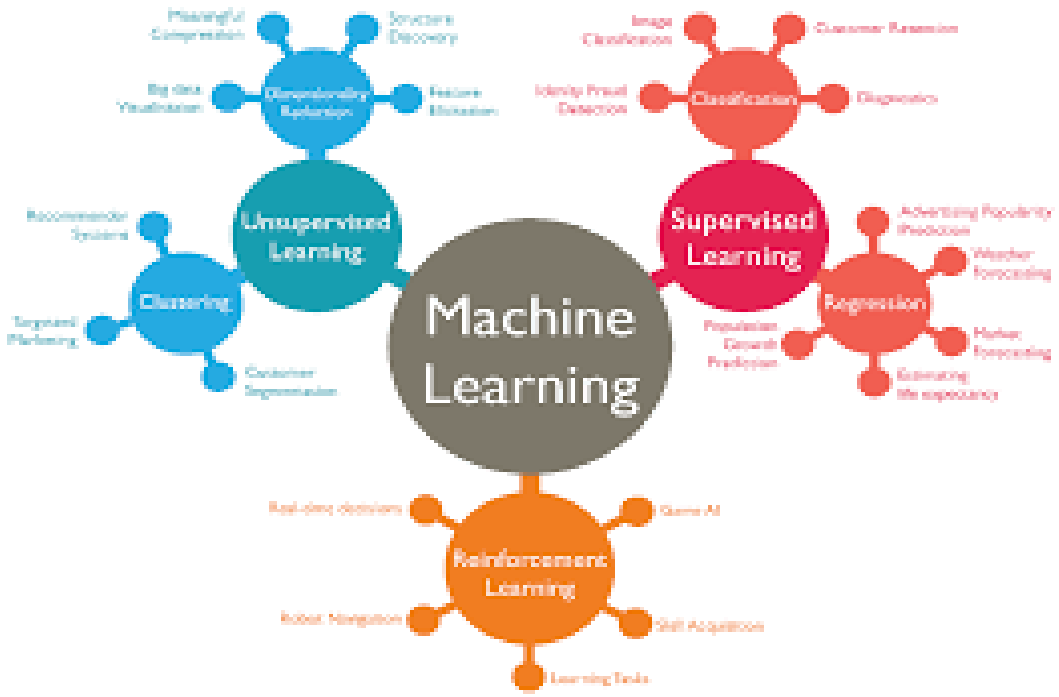
Challenges in spam filtering:

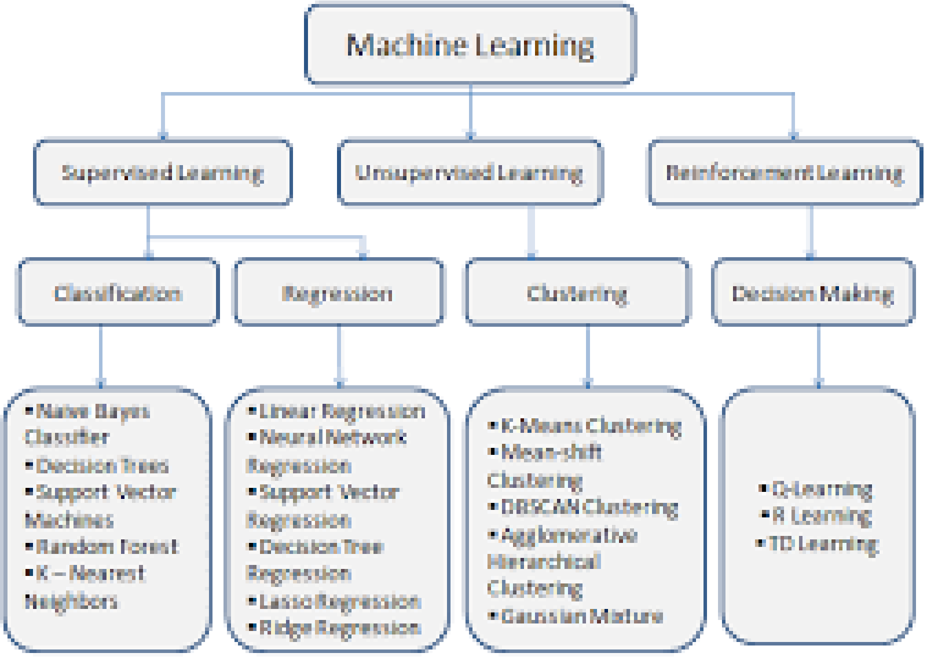
There are even dedicated academic conferences, talking about new ideas in spam filtering, such as CEAS and the MIT Spam Conference.

The first automatic spam filters were probably the USENET cancelbots, which used mathematical techniques such as the Breidbart Index (incidentally, we just passed the 16th anniversary on April 12 of the first major USENET spam). Since then, more and more techniques have been invented -- some have fallen into disuse, but others have found their way into common usage in some kind of Darwinian fight for the survival of the fittest techniques.

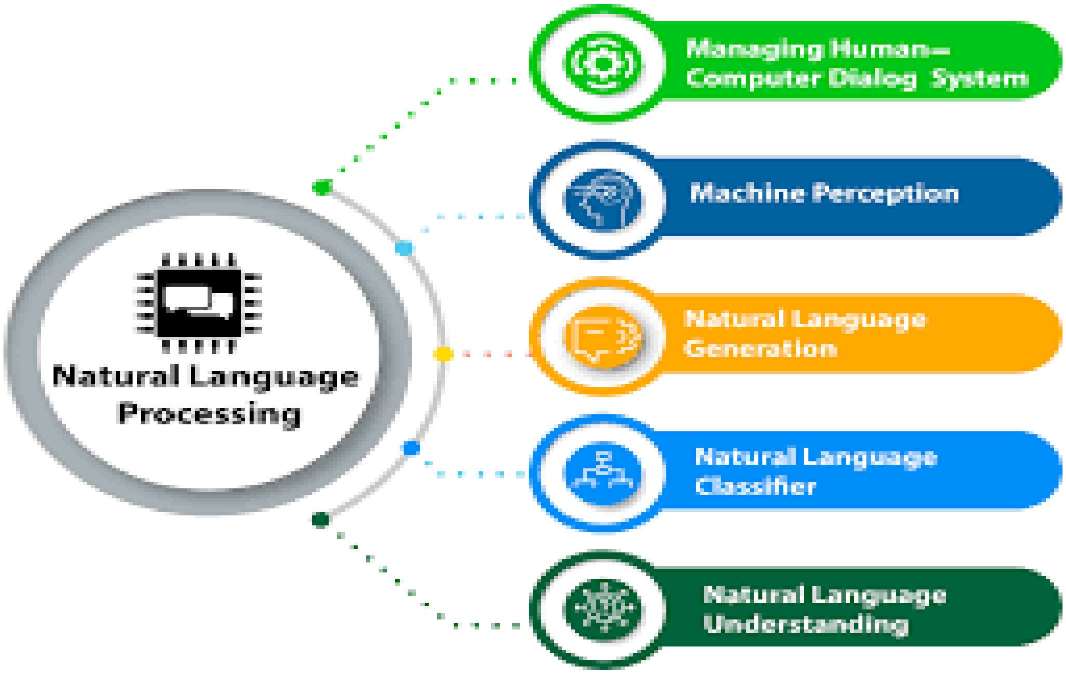
Machine learning:

Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.

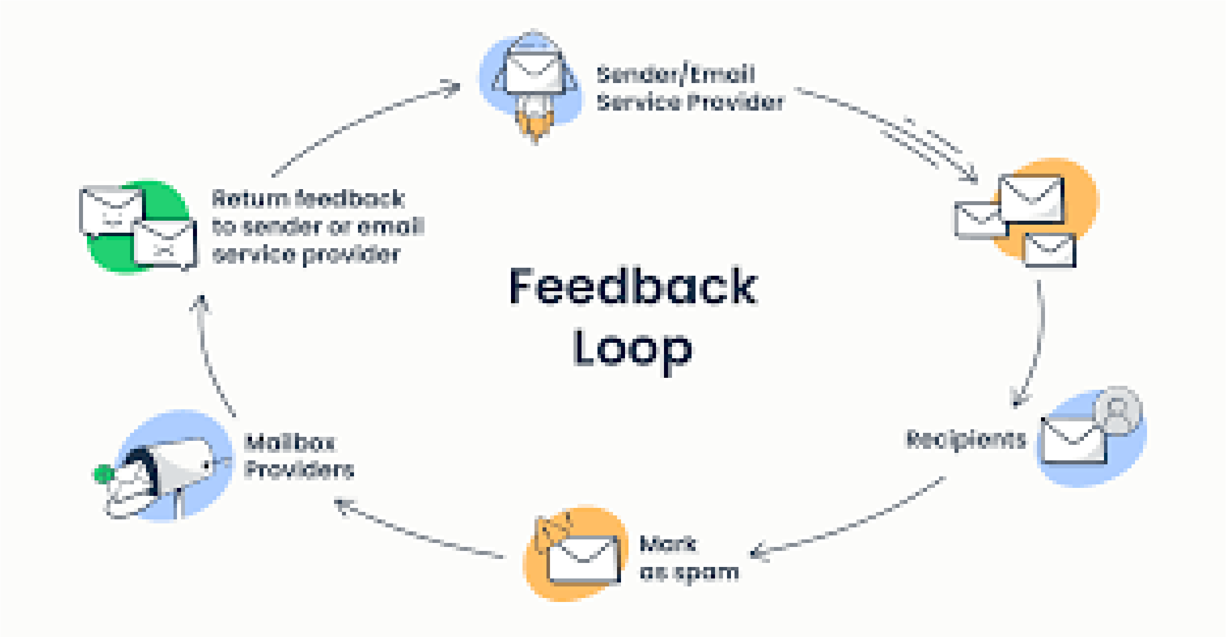




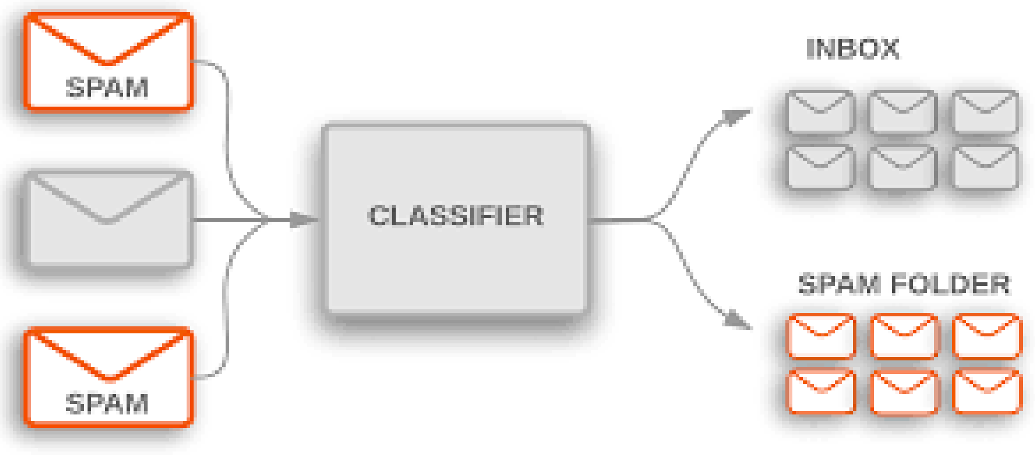
Natural language processing:


Natural Language Processing
Managing Human–Computer Dialog System
Machine Perception
Natural Language Generation
Natural Language Classifier
Natural Language Understanding

Feedback loop:


Feedback Loop
Sender/Email Service Provider
Recipients
Mark as spam
Mailbox Providers
Return feedback to sender or email service provider

STEPS INVOLVED IN CLASSIFICATIOON OF SPAM:



TEXT CLASSIFICTIONS:

1.Deep Learning and Neural Networks:

Deep learning is a subset of machine learning that utilizes neural networks with multiple layers to model and solve complex tasks, mimicking the human brain's structure

2.Transformer-Based Models:

Transformer-based models, like the BERT and GPT architectures, have revolutionized natural language processing tasks by using self-attention mechanisms to capture contextual information efficiently.

3.Ensemble Methods:

Ensemble methods combine multiple machine learning models to improve overall predictive performance by leveraging the strengths of different models.

4.Transfer Learning:

Transfer learning involves using pre-trained models on one task to enhance the performance of a related task, reducing the need for extensive training data.

5.Active Learning:

Active learning is a strategy where a model interacts with a human annotator to select and label the most informative data points, reducing the labeling effort required.

6.Explainable AI (XAI):

Explainable AI focuses on making machine learning models and their decisions more interpretable and transparent, enabling users to understand why a model makes specific predictions

7.Multi-Modal Learning:

Multi-modal learning combines information from different data types, such as text, images, and audio, to improve model performance and understanding.

8.Privacy-Preserving Techniques:

These methods protect sensitive data during machine learning by applying encryption, anonymization, or differential privacy.

9.Cross-Lingual Spam Detection:

Techniques for detecting and filtering spam content in multiple languages, often using NLP and machine learning methods.

10.Behavioral Analysis

This involves studying patterns in human behavior, often with the assistance of machine learning models, to gain insights or detect anomalies.

11.Graph-Based Approaches:

Graph-based models use network structures to represent and analyze relationships between data points, making them useful for tasks involving complex connections.

12.Natural Language Generation (NLG) Detection:

NLG detection aims to distinguish between human-generated and machine-generated text, often used to combat misinformation.