

BUILDING A SMARTER AI-POWERED SPAM CLASSIFIER

Project Overview:

In this part you will need to understand the problem statement and create a document on what have you understood and how will you proceed ahead with solving the problem. Please think on a design and present in form of a document.

Problem Definition:

The problem is to build an AI-powered spam classifier that can accurately distinguish between spam and non-spam messages in emails or text messages. The goal is to reduce the number of false positives (classifying legitimate messages as spam) and false negatives (missing actual spam messages) while achieving a high level of accuracy.

Design Thinking:

1.Data Collection: We will need a dataset containing labeled examples of spam and nonspam messages. We can use a Kaggle dataset for this purpose.

2.Data Preprocessing: The text data needs to be cleaned and preprocessed. This involves removing special characters, converting text to lowercase, and tokenizing the text into individual words.

3.Feature Extraction: We will convert the tokenized words into numerical features using techniques like TF-IDF (Term Frequency-Inverse Document Frequency).

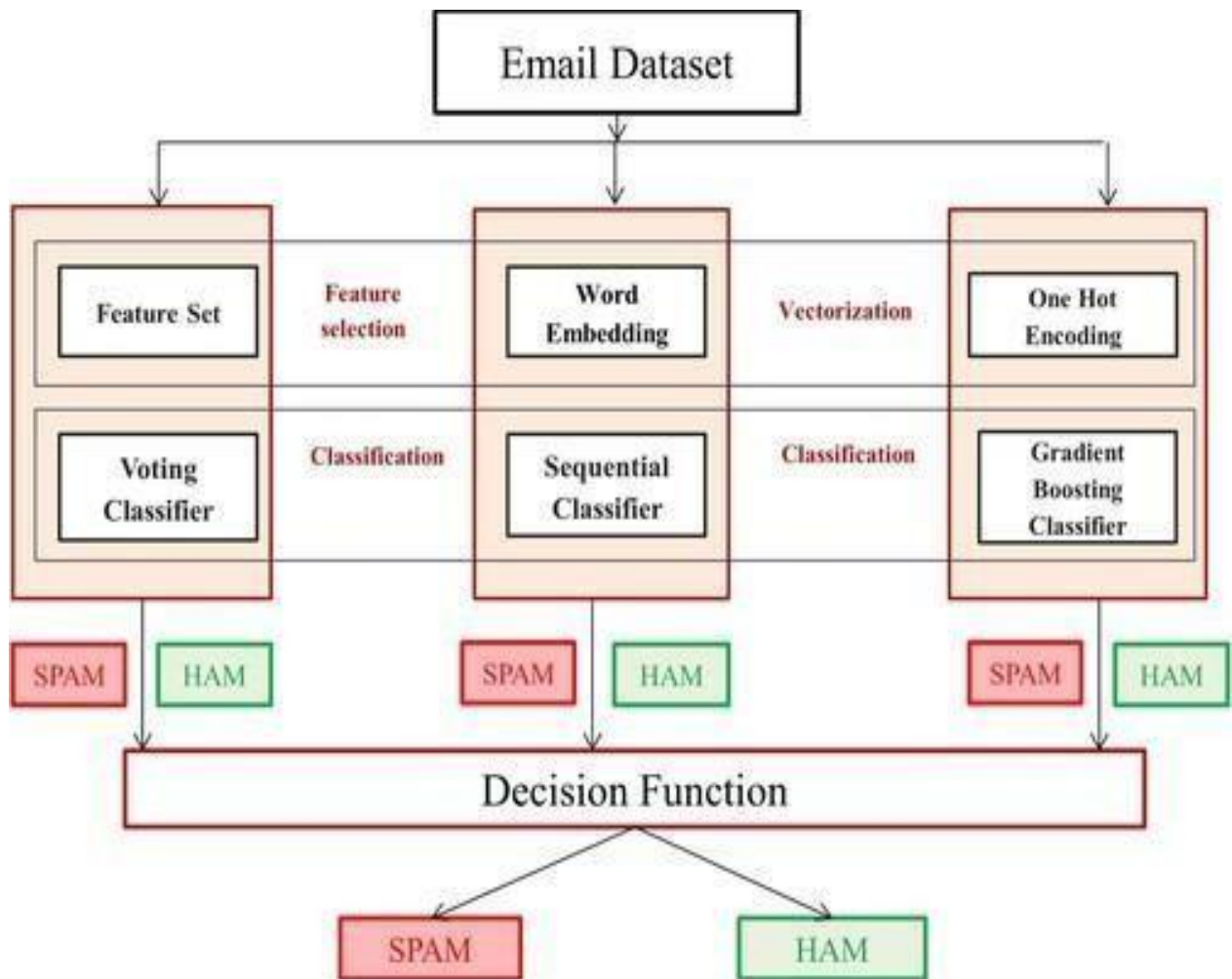
4.Model Selection: We can experiment with various machine learning algorithms such as Naive Bayes, Support Vector Machines, and more advanced techniques like deep learning using neural networks.

5.Evaluation: We will measure the model's performance using metrics like accuracy, precision, recall, and F1-score.

6.Iterative Improvement: We will fine-tune the model and experiment with hyperparameters to improve its accuracy.

Dataset Link: <https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset>

ARCHITECTURE FOR SPAM CLASSIFIER

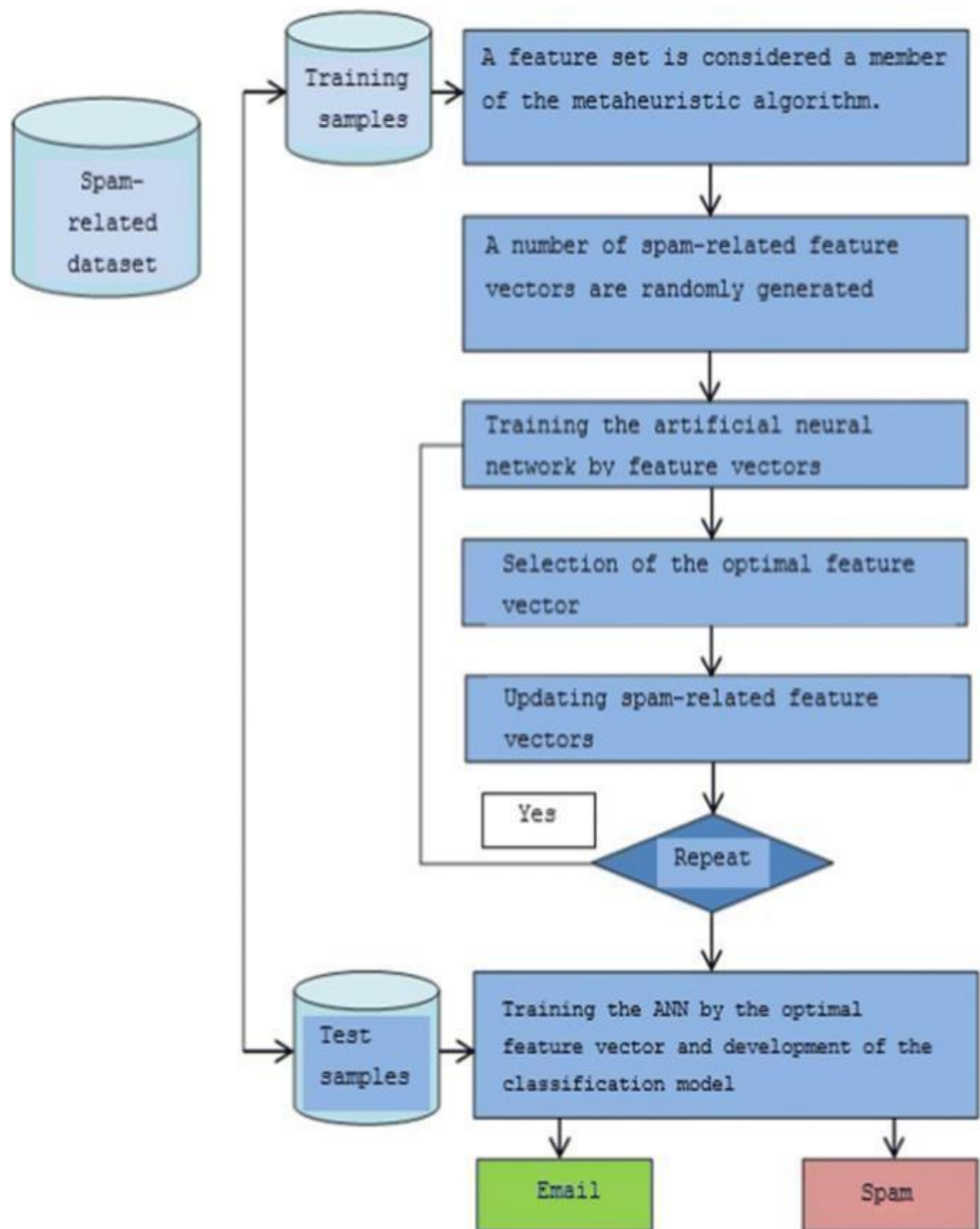


1. In this architecture we divide mail into two parts as SPAM and HAM.
2. We are using three types of classifiers to classify the mails following as:

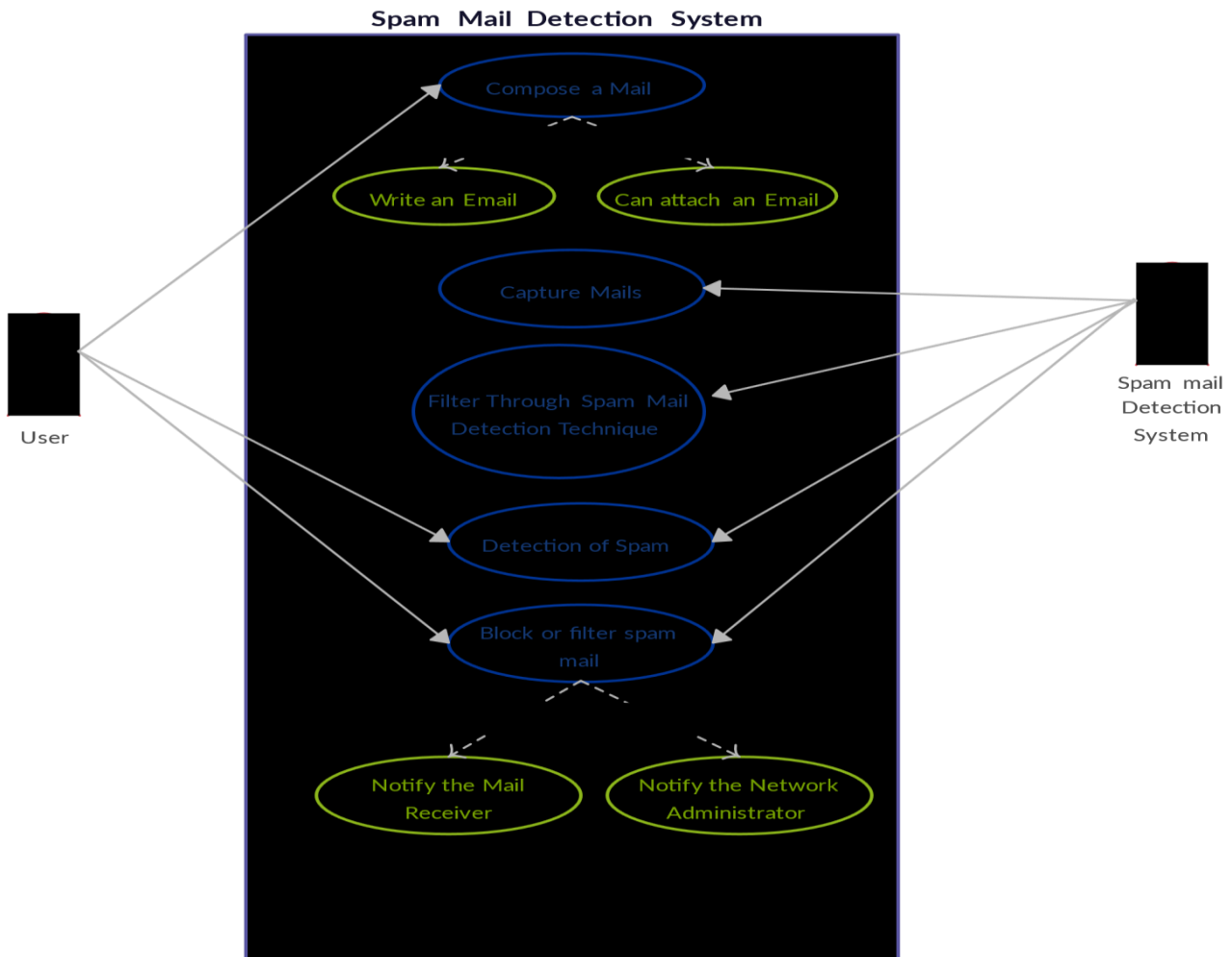
- Voting Classifier
- Sequential Classifier
- Gradient Bossting Classifier

3. These three classifiers are used to filter messages, if spam or not spam.

FLOWCHART



USER DIAGRAM FOR SPAM CLASSIFIER



1. Load and simplify the dataset.
2. Explore the dataset.
3. Handle the imbalance datasets.
4. Split the dataset.
5. Classify the dataset using algorithms like Random Forest, Naïve Bayes Etc.,
6. Finally find the messages spam or not spam.