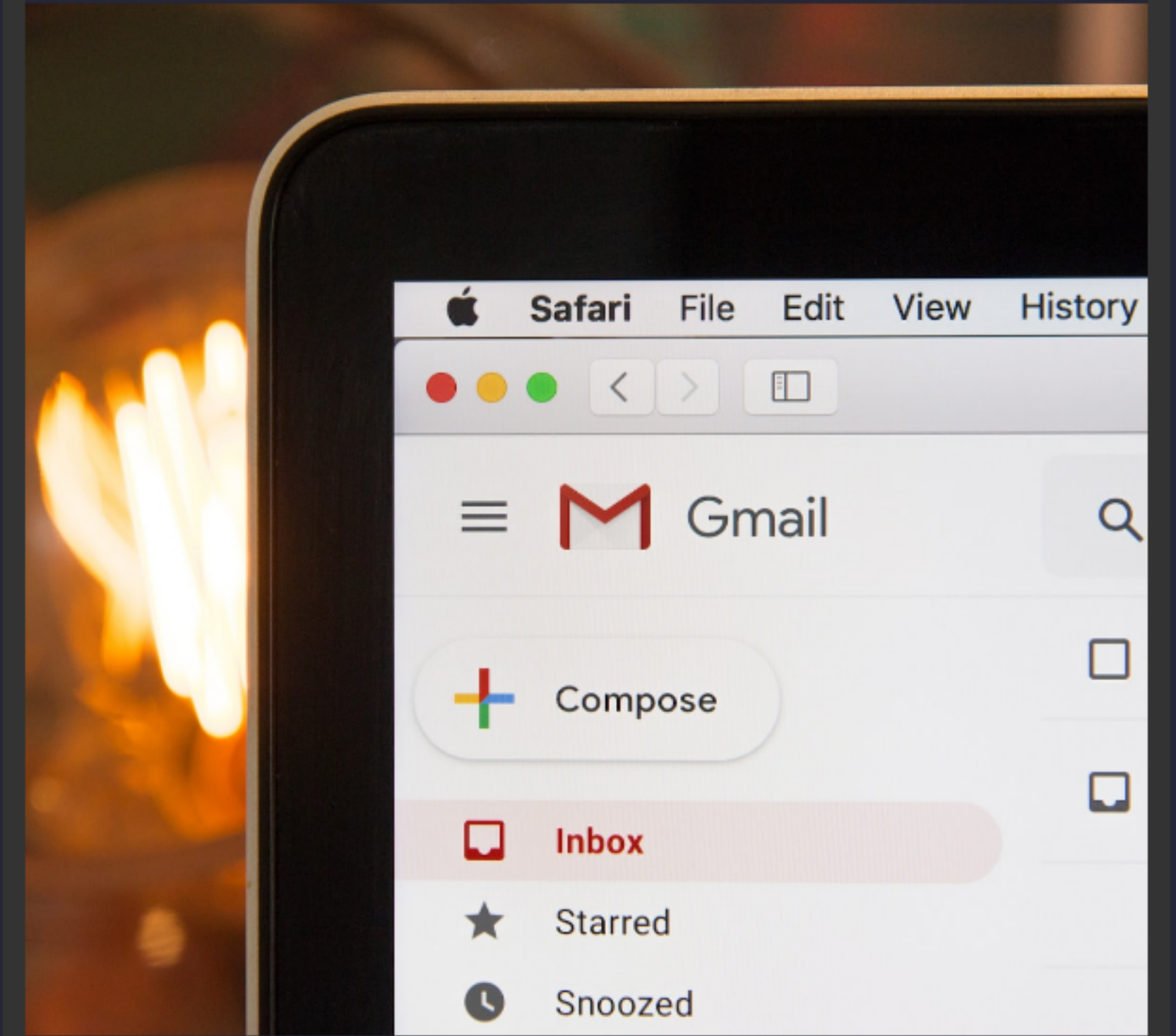


Detecting Spam Mail:

Separating the Ham from the Spam for
Mail Users



What is spam mail

Defining and explaining its characteristics

Spam mail is unsolicited and often malicious email that can contain unwanted advertisements, phishing attempts, or malware.



Packages and Models

- Numpy
- Pandas
- TfidfVectorizer
- Logistic Regression

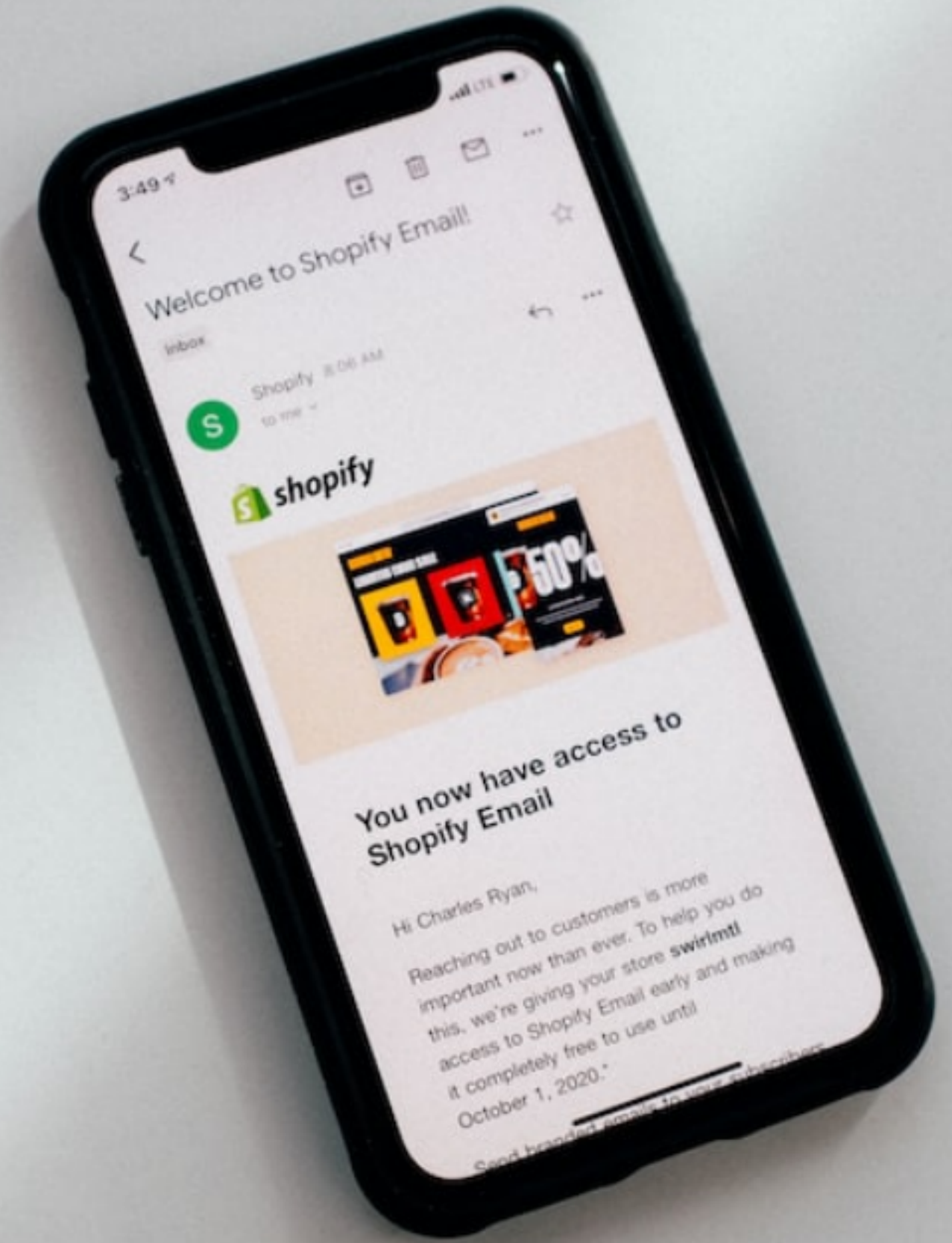
Collecting email data

The first step is to collect email data, which can be done by accessing email servers or by using email datasets available online



Feature extraction:

Once the data is preprocessed, we need to extract relevant features from it. This can be done using techniques like Bag of Words, TF-IDF, or word embeddings..



Blacklists and Whitelists

Detecting spam mail

Blacklists and whitelists are methods used to identify spam mail by either blocking or allowing email addresses, domains, or IP addresses based on predefined lists.

