

## Abstract

During the early days of email communication in the past, users were facing a growing problem: spam emails. Gmail launched in 2004. Emails became a primary method of communication, therefore the need for spam filters was paramount. Over time, Gmail has managed to block over 99% of spam reaching inboxes. Machine-based learning powered by user feedback is used to catch spam mail. (Kumaran, N. 2022). There were several legal, social, ethical, and professional issues to consider too, including the safeguarding of user data, allowing smooth and easy communication world-wide, and being able to balance effective spam filtering with privacy and confidentiality.

## The spam problem in Gmail:

Gmail was facing a big problem with spam emails. A substantial number of users' inboxes were being flooded with unwanted emails, scams, and inappropriate content. In December 2019, 85% of mail reaching users' inboxes was spam (Cloud, N. 2020). Users of Gmail can range from large businesses to personal accounts. This was a serious problem because it not only affected user experience, but it posed big security risks, financial scams and the spreading of viruses and malware. The enormous volume of spam emails was threatening the reliability and efficiency of Gmail. A solution needed to be found and this is where computer science comes in. Gmail introduced filters that help solve this problem. Their filters are built to give users a great overall experience when using Gmail. (Kumaran, N. 2022).

## The Gmail spam filters:

When Gmail first introduced spam filters, they focused on improving user engagement. Their spam filtering system used a combination of algorithms, machine learning and neural networks (Cloud, N. 2020).

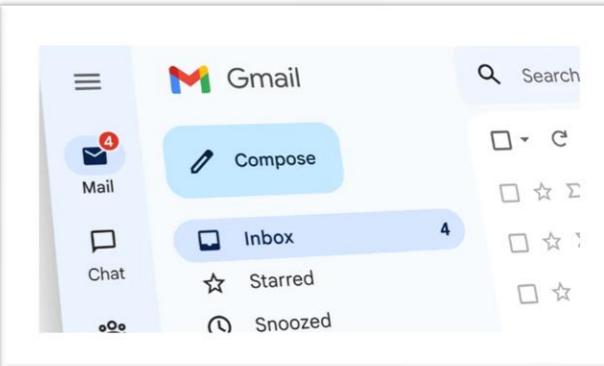
The key features of the Gmail spam filter include:

- 1) Filtering that is based on rules. In this early stage of filtering, users were able to establish specific rules and apply them to any incoming emails (Zola, A. 2021).
- 2) Machine learning (algorithms) powered by user feedback. Gmail uses AI-driven filters that determine what emails get marked as spam. These filters look for different signals including the IP address and domains/subdomains (Kumaran, N. 2022).
- 3) Neural networks – they have the ability to learn and identify spam emails by analysing a bulk load of emails (Gbenga, Da. et al. 2019). This method usually analyses words, how often they occur, how they are distributed and used in an email.
- 4) Gmail uses optical character recognition (OCR) to protect users from spam images/photos (Gbenga, Da. et al. 2019).
- 5) User feedback mechanisms is a key part to the Gmail spam filtering process. The filters learn from user actions.

# Gmail Spam Filters

Maryam Akhtar

Student Number: 239528371



### Social:

Being able to obtain faith/trust in users and maintaining it in digital communication. This can be done by preventing scams and inappropriate content.

Enabling users from all around the world to have a secure and easy way of communication.



### Legal:

Areas of legal compliance that need to be considered include data protection laws and knowing the regulations about safeguarding user data.

Having user consent – being clear and open about user data usage and getting consent from users to scan emails (Gbenga, Da. et al. 2019).

### User Education:

It is of paramount importance to educate users of spam/phishing and how they can avoid these types of emails. If users are educated on this, they will be able to use their email accounts effectively and keep cybersecurity in mind.

### Effectiveness of Gmail spam filters:

The Gmail spam filters are designed to detect and prevent unwanted emails from reaching users' inboxes. One reason why the spam filter used is so effective is because it works by using natural language processing to check incoming emails. Google stated that the spam filter was so effective that they managed to control 99% of spam from user inboxes. They added that their false positive rate (emails identified incorrectly as spam) was only 1%. (Kumaran, N. 2022).

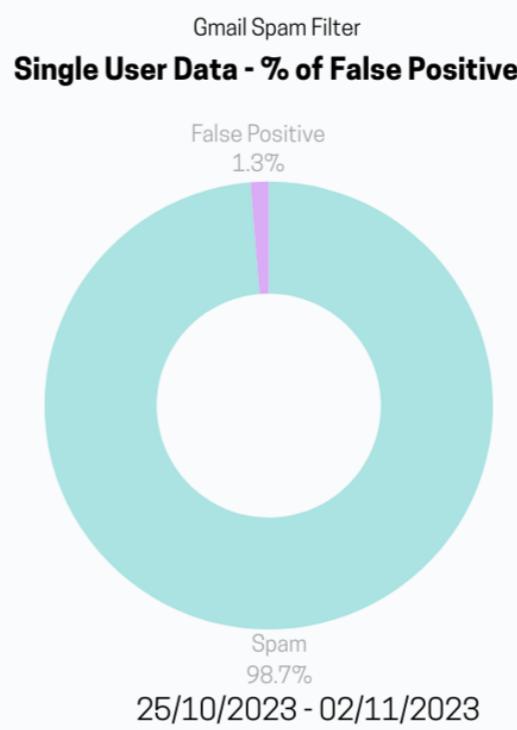
### Single-use data/metrics:

Using myself as an example to look at the effectiveness of the spam filters used by Gmail, I recorded data of emails I received in the past seven days.

This is what I found (pie chart showing data):

- I received a total of 159 emails over the week.
- Of the 159 emails, 152 of them went to my spam folder.
- Of the emails arriving at my inbox, none of them were spam.
- Of the emails in my spam folder, 2 of them were misidentified as spam.

Overall, as seen on the pie chart (on the right), 1.3% spam emails was a false positive. This shows that there is a very small percentage of error and during the past week of data, the Gmail spam filter was effective.



### Ethical:

Being fair and taking biases into consideration to ensure that any user discrimination is avoided as well as protecting the privacy of users and ensuring their data is kept confidential (Gbenga, Da. et al. 2019).

### Professional:

Ensuring that professionals within this field are receiving continuous professional development (CPD) training to stay up-to-date with changes/improvements in cybersecurity and laws.

Professionals in this area should address the challenges in cybersecurity and collaborate to keep users safe online.

### References:

Kumaran, N. (2022) *An overview of Gmail's Spam Filters* | Google Workspace Blog (2022) *Understanding Gmail's spam filters*. Available at: <https://workspace.google.com/blog/identity-and-security/an-overview-of-gmails-spam-filters> (Accessed: 01 December 2023).

Cloud, N. (2020) *Growth of Gmail Spam Filters* | Email Deliverability Perspective. Available at: <https://netcorecloud.com/blog/growth-of-gmail-spam-filters-email-deliverability-perspective/> (Accessed: 01 December 2023).

Zola, A. (2021) *What is a spam filter?*, Security. Available at: <https://www.techtarget.com/searchsecurity/definition/spam-filter#:~:text=Rule%2Dbased%20filters%20enable%20users,in%20the%20message%20or%20header>. (Accessed: 02 December 2023).

Gbenga, Da. et al. (2019) *Machine learning for email spam filtering: Review, approaches and open research problems*, *Heliyon*. Available at: <https://www.sciencedirect.com/science/article/pii/S2405844018353404> (Accessed: 02 December 2023).